

Überlagerte Funksignale in drahtlosen Sensornetzwerken

Von der Carl-Friedrich-Gauß-Fakultät für Mathematik und Informatik der
Technischen Universität Carolo-Wilhelmina zu Braunschweig genehmigte
Dissertation zur Erlangung des Grades eines Doktors der
Ingenieurwissenschaften (Dr.-Ing.)

von Dipl.-Ing. Albert Krohn

Referent: Prof. Dr. M. Beigl, Institut für Betriebssysteme
und Rechnerverbund, Technische Universität
Braunschweig

Korreferent: Prof. Dr. W. Juling, Institut für Telematik,
Universität Karlsruhe (TH)

Datum der Promotion: 9. März 2007

Danksagung

Die vorliegende Dissertation entstand im Rahmen meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Institut für Telematik der Universität Karlsruhe (TH). Ich trat dort im September 2002 in die Forschungsgruppe des Telecooperation Office (TecO) ein, die von Prof. Dr.-Ing. Michael Beigl geleitet wurde. Das TecO war mir bereits aus meiner früheren Tätigkeit als Hilfwissenschaftler bekannt. Während dieser Zeit wurde mein Interesse für die Forschung dieser Gruppe im Bereich Ubiquitous Computing geweckt und motivierte mich zu einem Promotionsvorhaben. Die flachen Hierarchien und selbstverantwortliche Arbeit in TecO schufen während meiner Zeit als Student und insbesondere während meiner Tätigkeit als wissenschaftlicher Mitarbeiter eine besonders positive Arbeitsatmosphäre. Der tägliche Austausch mit den Mitarbeitern und Studenten, sowie deren breites Interesse an vielen Forschungsfeldern erzeugte eine kreative, motivierende und stimulierende Umgebung, in der das Arbeiten und Forschen viel Freude gemacht hat. Für die Forschung an aktuellen Themen ist heute auch ein internationaler Austausch unabdingbar. Das TecO hat mir die Möglichkeit gegeben, während meiner Zeit als Mitarbeiter eine Vielzahl von Konferenzen und Partnerinstituten in Deutschland und im Ausland zu besuchen. Die Bekanntschaft und der Austausch mit den Forschern in gleichen oder ähnlichen Wissenschaftsbereichen beurteile ich im Nachhinein als eine der wichtigsten Erfahrungen und als wesentlichen Antrieb für meine eigene Arbeit. Nicht zuletzt sind so auch Freundschaften über die professionelle Zusammenarbeit hinaus entstanden.

Besonderen Dank möchte hier Prof. Dr. W. Juling aussprechen, der als Leiter des Instituts gerne an unseren Vortragsreihen teilgenommen hat und mir bei meiner Promotion interessante Anregungen und nützliche fachliche Kritik gegeben hat und sich bereit erklärt hat, meine Arbeit als Korreferent zu betreuen. Neben den fachlichen Diskussionen hat auch seine langjährige Erfahrung mit Doktorarbeiten mir sehr geholfen, meine wissenschaftlichen Beiträge klarer darzustellen. Weiterhin möchte ich Prof. Dr. rer. nat. H. Hartenstein meinen Dank aussprechen, der mit seiner fachlichen Kompetenz besonders in der Modellbildung der Theorie hilfreiche Kritik geleistet hat. Darüber hinaus möchte ich Prof. em. Dr. Dr.-Ing. E.h Dr. h.c. mult. Gerhard Krüger dafür danken, dass er immer wieder Zeit für Be-

suche im TecO gefunden und besonders die Strukturierung und Argumentation in meiner Arbeit in die richtigen Wege geleitet hat.

Den Mitarbeitern am TecO bin ich zu außerordentlichem Dank verpflichtet. Die vielen Diskussionen und gemeinsamen Veröffentlichungen mit Tobias Zimmer und Christian Decker und der Austausch über weit gefächerte Themen waren inspirierend und haben mir geholfen, meine eigenen Ideen besser zu formulieren und wichtige Aspekte aufzudecken. Weiterhin möchte ich dem Leiter der TecO-Gruppe, Prof. Dr.-Ing. Michael Beigl für seinen außerordentlichen Beitrag zum Gelingen meiner Arbeit danken. Er hat sich stets für meine Ideen und Vorschläge interessiert und war ein wichtiger Diskussionspartner. Unter seiner professionellen Betreuung konnte ich meine Beiträge in anerkannte wissenschaftliche Fachpublikationsforen einbringen, und damit bereits eine Position in der internationalen Forschung einnehmen. Er hat mich immer wieder ermutigt, meine Ideen zu Ende zu bringen und mir geholfen, die neuen Beiträge systematisch und im wissenschaftlichen Kontext selbst zu überprüfen und verständlich darzustellen.

Weiterhin gilt mein Dank Martin Bäuml und Sabin Wendhack, die durch ihre überdurchschnittlichen Leistungen als Hilfwissenschaftler am TecO die Implementierung und Evaluierung meiner Forschungsergebnisse ermöglicht haben.

Zuletzt möchte ich noch meinen Eltern danken, die mich in meiner Ausbildung an Schule und Universität stets unterstützt haben und auch mein Promotionsvorhaben ausdrücklich befürwortet haben.

Zusammenfassung

Die Miniaturisierung von Elektronik und elektronischen Geräten nimmt heute immer weiter zu. Sie hat vor allem zwei Nebeneffekte mit sich gebracht: Erstens sind durch die verringerte Größe auch automatisch die Kosten für elektronische Geräte gefallen; zweitens ist mit der geringen Größe auch der Energiebedarf deutlich kleiner geworden. Diese beiden Eigenschaften haben dazu geführt, dass Miniaturrechner (so genannte Mikroprozessoren) heutzutage in fast alle Bereiche der Technik Einzug genommen haben. Haushaltsgeräte, Unterhaltungselektronik und Verkehrstechnik sind ohne Mikroprozessoren heute nicht mehr denkbar. Viele dieser Geräte enthalten Sensorik und Benutzerschnittstellen, die von den Prozessoren gesteuert werden. Wenn Geräte mehr als nur einen Prozessor besitzen, ist es auch nötig, die Kommunikation zwischen den Prozessoren zu kontrollieren. Ein Forschungsbereich, der sich mit der drahtlosen Kommunikation solcher Mikroprozessoren beschäftigt, ist der der *Sensornetzwerke*. Die potentiell große Anzahl der teilnehmenden Geräte an solchen Netzwerken (sog. Sensorknoten), sowie deren Mobilität und Heterogenität erfordern neue Mechanismen, die über herkömmliche Kommunikationsprotokolle hinausgehen.

Die vorliegende Arbeit stellt ein neues Konzept der Anwendung von *überlagerten Funksignalen* vor, und zeigt ihren Nutzen besonders für Sensornetzwerke. Überlagerte Funksignale entstehen immer dann, wenn mehrere Kommunikationspartner im selben Funkkanal zur selben Zeit senden. Damit solche Überlagerung nicht im Chaos endet, trägt diese Arbeit zu allen heute ersichtlichen, relevanten Bereichen von überlagerten Funksignalen in Sensornetzwerken bei: Es werden Signalübertragung, Synchronisation, Kanalzugriff und Datenübertragung in den hier relevanten Zusammenhängen diskutiert. Die gewonnenen Erkenntnisse ermöglichen völlig neue Mechanismen, die speziell für mobile Szenarien mit extrem hohen Teilnehmeranzahlen geeignet sind. Ein Beispiel wäre hier das Protokoll SDJS (siehe Kapitel 7), das Datenübertragung auf Basis von überlagerten Funksignalen ermöglicht und dabei diese Überlagerung nutzt, um schon während der Übertragung die Daten zu verarbeiten. Damit wird eine *probabilistische* Datenkommunikation ermöglicht, die extrem schnell ist und mit einer sehr großen Anzahl von Teilnehmern arbeiten kann, ohne dass der Datendurchsatz darunter leidet. Um die Praxisrelevanz der Arbeit zu unterstreichen, wurden sämtliche neuen Mechanis-

men und Protokolle auf einer echten Sensornetzplattform, den Particle Computern, in Hard- und Software implementiert und quantitativ evaluiert.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Drahtlose Sensornetzwerke	2
1.2	Eigenschaften der Kommunikationszenarien	3
1.3	Einleitungsbeispiel für überlagerte Funksignale	4
1.4	Problemstellung und Motivation	5
1.5	Überlagerte Funksignale	6
1.6	These und Aufbau der Arbeit	7
2	Analyse Zielplattform und Kommunikation	10
2.1	Die Zielplattform	15
2.2	Bekannte Verfahren für überlagerte Funksignale	18
2.2.1	Opportunistic Large Arrays	19
2.2.2	Spectrum Pooling	22
2.2.3	dFSK	24
2.2.4	Kohärente Überlagerung	25
2.2.5	Kohärente Überlagerung zweier mobiler Quellen	27
2.2.6	Beam-forming mit Hilfe von kohärent überlagerten Funk- signalen	27
2.2.7	Virtuelle Radarbilder mit überlagerten Funksignalen	29
2.3	Bekannte Verfahren zur Zeitsynchronisation in Sensornetzwerken	30
2.3.1	LTS	32
2.3.2	RBS	32
2.3.3	Mini-sync und Tiny-sync	33
2.3.4	Überlagerte Funksignale zur Zeitsynchronisation	35
2.4	Thematische Einordnung und Positionierung dieser Arbeit	38
2.5	Zentrale Beiträge	41
2.6	Systemaufbau und Architektur	43
3	Übertragungsmodelle	45
3.1	Kooperatives Senden	52
3.1.1	Signalstatistiken für nicht kohärente Überlagerung	52

3.1.2	Statistik im Empfänger	54
3.1.3	Das Problem der Schwebungen	56
3.2	Verteiltes Senden	59
3.3	Gemeinsames Modell für ressourcenarme Knoten	60
3.4	Detektion des Rayleighparameters	61
3.5	ESK	62
3.6	Energienormierung	67
3.7	Zusammenfassung	69
4	Implementierung	70
4.1	Wie sehen im Bandpass überlagerte Funksignale tatsächlich aus? .	71
4.2	Verhalten sich überlagerte Signale und Leistungen tatsächlich additiv?	73
4.3	Kann man mit überlagerten Funksignalen das Sensor-Reach-back-Problem lösen?	75
4.4	Zusammenfassung	77
5	AwareCon: Einsatz von überlagerten Funksignalen zur Verbesserung von Netzwerkperformance	78
5.1	pPart Particle Computer	79
5.2	Das Protokoll <i>AwareCon</i>	80
5.3	Zeitsynchronisation in AwareCon	82
5.3.1	Aufbau der Synchronisation	84
5.3.2	Signalisierung	85
5.3.3	Evaluation der Synchronisation in AwareCon	87
5.3.4	Verteilte Synchronisation	89
5.4	Der Kanalzugriff in AwareCon	90
5.4.1	Überlagerte Funksignale als virtueller tri-state-Bus	92
5.4.2	Arbitrierung in AwareCon	94
5.4.3	TOMAC	94
5.5	Zusammenfassung	96
6	Erhöhung der Erreichbarkeit und Zuverlässigkeit mit überlagerten Signalen	98
6.1	Schlechte Abdeckung	100
6.2	Das Konzept von kooperativem Senden	101
6.2.1	Ausbreitungs- und Kanalmodell	102
6.2.2	Transportszenarien	103

6.2.3	Prinzipien der kooperativen Kommunikation	104
6.3	Erhöhung der Reichweite	110
6.3.1	Grenzverhalten	114
6.3.2	Energieverbrauch und Implementierungsaufwand	116
6.4	Zusammenfassung	117
6.5	Bestehende Probleme für die Implementierung und weitere Verbesserungen	118
7	Cross-Layer Parameterschätzung mit überlagerten Funksignalen	121
7.1	Synchrone Jam Signale - SDJS	124
7.2	Parameterschätzung mit SDJS: Anzahl der Geräte	126
7.2.1	Analytisches Modell: Das Jäger-Enten-Problem	127
7.2.2	Maximum-Likelihood-Schätzer	130
7.2.3	Implementierung und Evaluierung	132
7.2.4	Experiment und Verifikation	133
7.2.5	Vorteile der Parameterschätzung mit SDJS	134
7.3	Parameterschätzung mit SDJS: Anwendung im Supermarkt	135
7.3.1	Einsatz von Informationstechnologie im Supermarkt	135
7.3.2	Studien im Supermarkt	136
7.3.3	Multi-SDJS für MHD-Profile auf RFID	137
7.3.4	Implementierung	143
7.3.5	Zusammenfassung	144
7.4	Multi-SDJS: Datenfusion von Temperaturmessungen	145
7.5	Parameterschätzung mit SDJS: Datenfusion für Lokationssysteme	147
7.5.1	Andere Arbeiten im Bereich Lokalisierung von Sensor-knoten	148
7.5.2	Anwendungen und Motivation	148
7.5.3	Methodik und Fehlermodelle	149
7.5.4	Fehlerkomponenten in Messungen	150
7.5.5	Die Plattform der Experimente	153
7.5.6	Empirische Fehlermodelle	156
7.5.7	Messen und Mittelung von Messwerten in Lokationssystemen	161
7.5.8	Mittelung über mehrere Messpfade	163
7.5.9	Datenfusion mit Multi-SDJS	167
7.5.10	Der Ablauf der Lokalisierung	168
7.5.11	Anwendung und Simulation	169
7.5.12	Zusammenfassung	175

8 Zusammenfassung und Ausblick	176
8.1 Die Zielplattform	176
8.2 These und Beiträge	177
8.3 Ausblick	180
Literaturverzeichnis	182
Abbildungsverzeichnis	192
Tabellenverzeichnis	197
Glossar	198
Anhang	
A Signalkonstellation	199
B Das AwareCon-Protokoll: Flussdiagramm	201

1 Einleitung

Die Miniaturisierung von Elektronik nimmt heute immer weiter zu. Gordon Moore — Mitbegründer von Intel — hat dies schon 1965 versucht zu quantifizieren. Er sagte voraus, dass die Zahl der Transistoren auf einem Chip sich alle 18 Monate verdoppeln würde. Er erwartete also eine stetige Verkleinerung der Strukturen. Bis zum heutigen Tag hat sich diese Voraussage in etwa erfüllt. Aktuelle Analysen gehen jedoch davon aus, dass diese Geschwindigkeit durch Quanteneffekte in den Strukturen bald an Grenzen stoßen wird. Andere Technologievisionen erwarten eine Ablösung der Halbleitertechnik durch alternative Technologien. Die Miniaturisierung von Elektronik erzeugt zwei für diese Arbeit wichtige Nebeneffekte: Die *Kosten* und der *Energiebedarf* von elektronischen Schaltungen haben rapide abgenommen. So waren vor 50 Jahren ganze Räume mit Röhren und Relais gefüllt, um eine Funktionalität zu erfüllen, die heute jeder Taschenrechner beherrscht, der nur mit einer kleinen Knopfzellenbatterie gespeist wird. Die geringen Kosten für integrierte Schaltungen hat in den letzten Jahren den Einzug von Mikroprozessoren in fast alle technischen Bereiche ermöglicht. In Autos findet man heutzutage durchschnittlich schon mehr als 50 Mikroprozessoren, die Motorsteuerungen, Klimatisierung, Unterhaltung und Sicherheitsaufgaben wahrnehmen.

Ende der 1980er Jahre wurde ein Wissenschaftszweig zum Leben erweckt, der dieses Szenario von vielen verteilten Computern auf die Spitze treibt. Ein viel zitierter Artikel, der 1991 von Mark Weiser im Scientific American veröffentlicht wurde [82], skizziert die Vision von *Ubiquitous Computing*: Wir sind umgeben von einer Vielzahl von Computern, die nicht nur in Autos, sondern in Alltagsgegenständen eingebettet sind und untereinander und mit ihren Benutzern kommunizieren. Diese ursprüngliche Idee wurde von da an in vielen Forschungsprojekten aufgegriffen und mit anderen Forschungsrichtungen verbunden. Die Relevanz dieser Idee wurde nicht zuletzt mit der Einführung einer Unzahl von neuen Begriffen unterstrichen, die alle auf dieser Idee fußen. Beispiele für diese Begriffe sind *Pervasive Computing* und *Ambient Intelligence*. Viele der wissenschaftlichen Beiträge in Pervasive Computing oder Ambient Intelligence befassen sich mit der expliziten oder impliziten Interaktion der Benutzer mit der neuen, durchdringen-

den und allgegenwärtigen Welt der Computer. Ein anderer Wissenschaftszweig, genannt *wireless sensor network* oder *drahtlose Sensornetzwerke* befasst sich mehrheitlich mit der technischen Darstellung von solchen Systemen und deren internen technischen Strukturen. Drahtlose Sensornetzwerke verfolgen die gleiche Grundidee wie Ubiquitous Computing. Die Forschung in diesem Bereich fokussiert sich allerdings hauptsächlich auf Fragestellungen der Kommunikation und Energieversorgung.

1.1 Drahtlose Sensornetzwerke

Drahtlose Sensornetzwerke sind Systeme, die aus einer Vielzahl von drahtlos miteinander kommunizierenden Kleinstrechnern zusammengesetzt sind. Diese Kleinstrechner übernehmen z.B. Messaufgaben (z.B. von physikalischen Parametern) und kommunizieren sie über eine Funkschnittstelle. Diese kleinste Einheit der Sensornetzwerke wird auch *Sensorknoten* oder *Funkknoden* genannt. In Abbildung 1.1 ist ein solcher Funkknoden¹ abgebildet. Die Motivation zur Erforschung solcher Netzwerke stammt ursprünglich aus dem militärischen Umfeld. Die ersten Systeme wurden für Überwachungsaufgaben während des kalten Krieges entworfen. Die moderne Forschung von Sensornetzwerken begann dann in den 1980ern. Die erste treibende Kraft war die Forschungseinrichtung der US-Streitkräfte, die DARPA (Defense Advanced Research Projects Agency). Mit ihrem „Distributed Sensor Network“ (DSN)-Programm startete sie die ersten Anstrengungen, diese verteilten Systeme genauer zu untersuchen.

Mit drahtlosen Sensornetzwerken können Messwerte wie z.B. Umweltparameter großräumig beobachtet werden. Dafür wird eine Vielzahl von Sensorknoten, die sowohl Sensoren als auch drahtlose Kommunikation beinhalten, auf das Beobachtungsgebiet ausgebracht. Die Sensorknoten besitzen in den meisten Fällen eine autarke Energieversorgung. Diese wird meist aus endlichen Ressourcen (z.B. Batterien) und regenerativen Ressourcen (z.B. Solarzellen) zusammengesetzt. Energie mit Hilfe von Solarzellen, Wärmetauschern und anderen Hilfsmitteln aus der Umwelt zu gewinnen, wird allgemein unter dem Begriff *energy harvesting* zusammengefasst.

Die Messwerte, die mit Hilfe solcher Sensorknoten erfasst werden, können über die Funkverbindungen an Datenverarbeitungssysteme angeschlossen wer-

¹Baugruppe „uPart“ der Firma Particle Computer GmbH [4]

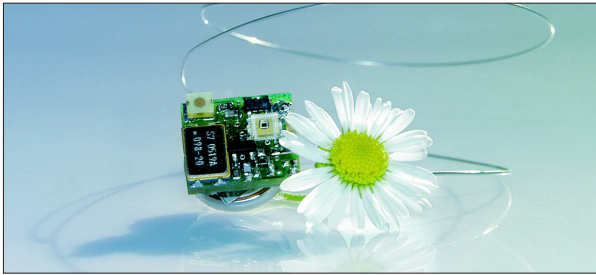


Abbildung 1.1. Ein Funkknoten ist die kleinste Einheit eines Sensornetzwerkes, hier im Vergleich zur Blüte eines Gänseblümchens

den. Ist die Entfernung vom Ort der Messung zur Zielverarbeitung zu weit, so wird in einem Sensornetzwerk eine Nachricht stückweise von Sensorknoten zu Sensorknoten weitergereicht, bis sie zum Ziel gelangt ist. Ein solches Weiterleiten von Nachrichten wird in drahtlosen Sensornetzwerken als *Multihop-Kommunikation* bezeichnet. Welchen Weg eine Nachricht durch ein solches Sensornetz nehmen sollte, kann aus unterschiedlicher Sicht auf das System nicht eindeutig beurteilt werden. Verfahren, die sich mit solcher Weiterleitung von Nachrichten in einem Sensornetzwerk — oder allgemein in einem Funknetzwerk — befassen, nennt man *Routing-Algorithmen*.

1.2 Eigenschaften der Kommunikationsszenarien

Kommunikation in Szenarien von Ubiquitous Computing und Sensornetzwerken unterscheidet sich in vieler Hinsicht von klassischer Funktechnik wie Richtfunk oder zellulärer Mobilfunk. Als erstes Unterscheidungsmerkmal kann die hohe *Anzahl* an kommunizierenden Geräten und ihre *Mobilität* betrachtet werden. Szenarien allgegenwärtiger Informationstechnologie und auch Sensornetzwerke zielen auf Systeme ab, die Anzahlen von über 1000 Endgeräten umfassen, die auf dichtem Raum zusammenarbeiten sollen. In visionären Szenarien sind z.B. Funkknoten in Alltagsgegenständen eingebettet oder werden am Körper oder sogar als Implantat getragen. Hohe Anzahlen von Endgeräten und Mobilität sind Eigenschaften, von denen der zelluläre Mobilfunk auch betroffen ist; dennoch wirken sich diese Eigenschaften auf Sensornetzwerke anders aus: Mobilität bedeutet hier

z.B., dass Verbindungen häufig abgebrochen werden und neue Verbindungen aufgebaut werden müssen. Beim zellularen Mobilfunk hingegen sind die Zellengrößen mit mehreren Kilometern so groß, dass Mobilität nur sehr gelegentlich zu einem Ausfall der Verbindung und deren Neuaufbau führt. Auch sind die Netzwerktopologien völlig unterschiedlich. Während zellulärer Mobilfunk auf eine Sternstruktur mit der Basisstation im Zentrum setzt, wird bei Sensornetzwerken auf diese zentrale Einheit verzichtet und alle Kommunikationspartner sind gleichberechtigt. Die starke und vielfältige Interaktion mit dem Benutzer führt auch zu anderen Arten von Datenverkehrsaufkommen. Ereignisbasierte Übertragung kann aufgrund der hohen Endgerätezahlen z.B. zu einer plötzlichen Explosion von Kanalzugriffen führen.

1.3 Einleitungsbeispiel für überlagerte Funkssignale

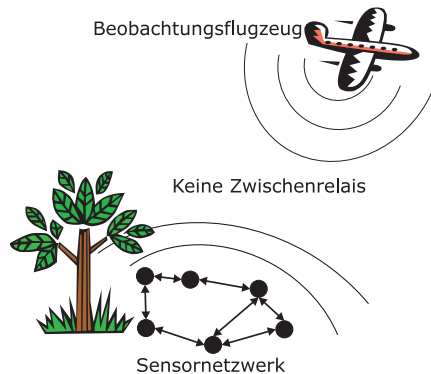


Abbildung 1.2. Das Sensor-Reach-Back-Problem

Als Einleitungsbeispiel für diese Arbeit soll ein Szenario aus der Welt der Sensornetzwerke dienen. In diesem Beispiel wird gezeigt, dass besonders die große Anzahl von Endgeräten im gemeinsamen Betrieb einen zusätzlichen Nutzen erzeugen kann, auch wenn das einzelne Gerät keine große Leistungsfähigkeit hat. Das Beispiel ist in Abbildung 1.2 dargestellt und bekannt als das

Sensor-Reach-Back-Problem. Ein Sensornetzwerk ist in einem Wald ausgebracht, um dort z.B. als Frühwarnsystem für Waldbrände zu dienen. Die

Sensorknoten sind also im Wald verstreut und überwachen z.B. die Temperatur, um Feuer zu entdecken. Ein Beobachtungsflugzeug fliegt über den Wald, um die Sensorwerte des Sensornetzes einzusammeln. Die Sensorknoten sind auf dem Waldboden nur wenige Meter voneinander entfernt und so ist es durch *Multihop*-Kommunikation möglich, dass sie untereinander Daten austauschen. Die Entfernung zum Flugzeug ist allerdings vergleichsweise hoch, so dass die Sendeleistung eines einzelnen Knotens nicht ausreicht, um seine Messwerte an das Flugzeug zu kommunizieren. Das Flugzeug ist jedoch in der Lage, die Sendeleistung aufzubringen, um Nachrichten an die Sensorknoten zu verschicken.

Für die Lösung könnte man hier z.B. eine Basisstation im Wald einrichten, die Kontakt zu dem Sensornetzwerk halten kann und über genügend hohe Sendeleistung verfügt. Eine solche Lösung ist in vielen Szenarien nicht sinnvoll, da man so die gesamte Kommunikation zum Sensornetzwerk über diese eine Basisstation führen müsste. Das erzeugt ein zusätzliches Ausfallrisiko. In Kapitel 6 wird diese Diskussion noch detailliert weiter geführt werden. Für das Einleitungsbeispiel soll es jedoch genügen, dass nun zum Systemdesign eine Lösung über eine Basisstation nicht verfolgt werden soll. Stattdessen soll versucht werden, die Sendeleistungen der Sensorknoten zu *akkumulieren*, um so eine gemeinsame, höhere Sendeleistung zu erreichen. Dafür einigen sich die Sensorknoten auf eine Nachricht, die sie dann gemeinsam und zur gleichen Zeit verschicken würden, um so die Sendeleistung und damit die Reichweite zu erhöhen. Diese naheliegende Idee führt nun direkt zum Thema der vorliegenden Arbeit: *Überlagerte Funksignale für Sensornetzwerke*. Das zeitgleiche Versenden einer Nachricht aus verschiedenen Quellen impliziert ja gerade die Überlagerung der Signale auf dem Medium. Überlagerte Funksignale scheinen eine brauchbare Lösung zu sein, um das Einleitungsbeispiel technisch zu lösen. Mit überlagerten Funksignalen können jedoch noch weitere, grundsätzliche Probleme in drahtlosen Sensorsystemen gelöst werden.

1.4 Problemstellung und Motivation

In allen Ebenen der Kommunikation in den Szenarien von Ubiquitous Computing und Sensornetzwerken spielt *Skalierung* eine entscheidende Rolle. Die große Anzahl von Endgeräten stellt neue Herausforderungen an die Protokolle und Me-

chanismen in allen Netzwerkschichten dar. Die vorliegende Arbeit leistet insbesondere auf *vier* Problembereichen in Sensornetzwerken Beiträge:

Zeitsynchronisation: Sensorknoten müssen für viele Anwendungen und Funktionalitäten gut synchronisierte, lokale Uhren haben. Dies ist z.B. wichtig, um durch zyklische Schlafzeiten der Netzwerkteilnehmer Energie zu sparen.

Kanalzugriff: Durch die Miniaturisierung kann schon heute eine Vielzahl von Funkknoten in einer Funkzelle zusammen existieren. Es ist notwendig, die potentiell große Anzahl an Knoten, die zeitgleich auf den Kanal zugreifen, zu koordinieren. Herkömmliche Verfahren sind hier nicht ausreichend.

Zuverlässigkeit: Besonders bei der Überwachung von kritischen Messgrößen für z.B. Frühwarnsysteme spielt Zuverlässigkeit eine entscheidende Rolle. Es zeigt sich schon heute, dass in kommerzieller Anwendung besonders diejenigen Systeme Erfolg haben, die eine hohe Zuverlässigkeit erbringen können.

Datenfusion: Durch die Miniaturisierung und den geringen Preis eines einzelnen Sensorknotens werden Sensorsysteme oft mehrfach redundant ausgelegt. Die große Anzahl von Sensoren kann in den angeschlossenen Datenverarbeitungssystemen zu einer unbeherrschbaren Datenflut führen. Deshalb müssen Sensorsysteme Mechanismen anbieten, die bereits im Funknetz eine Vorverarbeitung der verteilt gemessenen Daten ermöglichen.

Für alle genannten Bereiche gibt es bereits Lösungen, die weitestgehend auf klassischer Netzwerktheorie aufbauen. Sie erreichen jedoch nicht die Effektivität und Effizienz, die für den realen Einsatz wünschenswert sind. Dieses Defizit motiviert neue Mechanismen zu erforschen, die Lösungen zu diesen speziellen Herausforderungen bieten können. In Abbildung 1.3 ist diese Motivation noch einmal dargestellt.

In der vorliegenden Arbeit werden die vier Problemstellungen mit einem einheitlichen, neuartigen Konzept adressiert. Es ist die Idee der *überlagerten Funksignale*.

1.5 Überlagerte Funksignale

Überlagerte Funksignale sind Signale, die zeitgleich im selben Frequenzbereich von verschiedenen Quellen ausgesendet werden und sich deshalb im Frequenz-

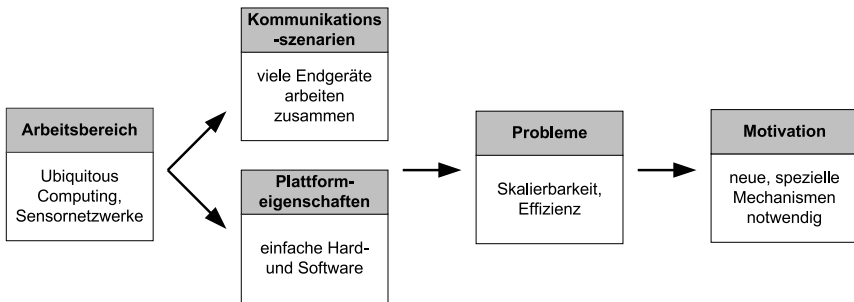


Abbildung 1.3. Die Probleme, die aus dem Arbeitsbereich resultieren, motivieren zu näheren wissenschaftlichen Betrachtungen

und Zeitbereich auf dem Übertragungsmedium überlagern. Mit überlagerten Funksignalen wird in dieser Arbeit ein neues Konzept vorgestellt, das speziell die vier oben genannten Probleme in Sensornetzwerken löst. Überlagerte Funksignale sind ein sehr junges Forschungsthema. Die ältesten Arbeiten, die im Zusammenhang mit Sensornetzwerken stehen, sind weniger als zehn Jahre alt. Die neuen Techniken überlagerter Funksignale ermöglichen speziell für Sensornetzwerke Mechanismen, die mit klassischer Herangehensweise nur schwer oder gar nicht zu erreichen sind. Dies liegt vor allem daran, dass Sensornetzwerke durch ihre Einschränkungen in Energie und Rechenleistung nur sehr einfache technische Lösungen unterstützen können. Überlagerte Funksignale können trotz dieser Einschränkungen viele neue Mechanismen ermöglichen, die die Grundprobleme in Sensornetzwerken lösen.

1.6 These und Aufbau der Arbeit

THESE:

In leistungsschwachen Sensornetzwerken lösen überlagerte Funksignale grundsätzliche Probleme wie Synchronisierung, Zuverlässigkeit, Kanalnutzung und Datenfusion.

Diese These wird im Laufe der vorliegenden Arbeit durch Analysen und Modelle unterstützt und wissenschaftlich nachgewiesen. Im Einzelnen werden genau

dort Beiträge vorgestellt, wo klassische Kommunikationstheorie keine geeigneten Lösungen für die neuartigen Kommunikationsszenarien bieten kann. Skalierung und Effizienz stehen dabei besonders im Vordergrund. Beispielhaft sei hier auf das Kapitel 7 verwiesen; dort wird ein Kommunikationsprotokoll vorgestellt, dessen Prozesszeit *unabhängig* von der Anzahl der Kommunikationsteilnehmer konstant ist. Es adressiert damit genau das Skalierungsproblem in Ubiquitous Computing und Sensornetzwerken. Darüber hinaus sind die Beiträge dieser Arbeit speziell auf sehr einfach aufgebaute Endgeräte ausgerichtet, denn ein massenhafter Einsatz von Informationstechnologie ist nur dann kostengünstig realisierbar, wenn die technische Komplexität eines einzelnen Gerätes stark beschränkt ist. Neuartige Mechanismen für die Kommunikation müssen dieses Leistungsdefizit der potentiellen Endgeräte unbedingt beachten. Die vorliegende Arbeit definiert den technischen Arbeitsbereich über die *Zielformat* in Kapitel 2.1. Diese ist besonders dadurch charakterisiert, dass sie sehr geringe Leistungsfähigkeit bezüglich z.B. Daten- und Signalverarbeitung hat. In Abbildung 1.4 ist der Aufbau der vor-

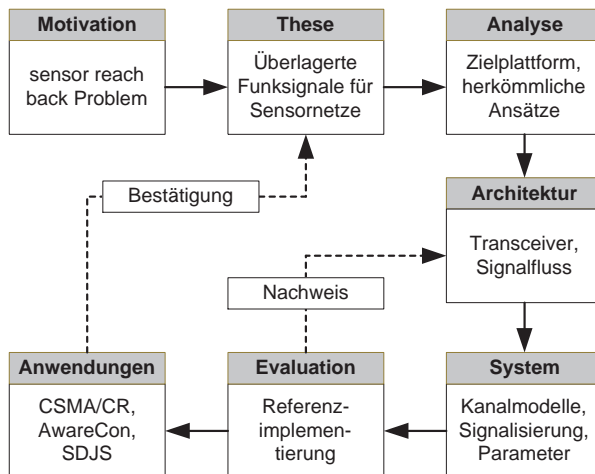


Abbildung 1.4. Aufbau der Arbeit

liegenden Arbeit dargestellt. Nach der Einleitung und der Hinführung zum Thema mit dem Einleitungsbeispiel umfasst Kapitel 2 eine genaue Rezension anderer Arbeiten auf dem gleichen Gebiet, sowie eine Analyse der drahtlosen Sensornetze

als Zielplattform. Aus dieser Analyse heraus werden Signalmodelle gewonnen und ein neues Modulationsverfahren in Kapitel 3.5 speziell für überlagerte Funksignale entworfen. In Kapitel 4 werden die theoretisch hergeleiteten Signalmodelle und Übertragungsmechanismen in einer echten Hardwareimplementierung bestätigt. Kapitel 5, 6 und 7 zeigen dann die vielseitigen Anwendungsmöglichkeiten von überlagerten Funksignalen in Sensornetzwerken. Dort werden die Probleme *Synchronisation*, *Zuverlässigkeit* und *Datenfusion* mit Hilfe von überlagerten Funksignalen gelöst. Die Beiträge dieser Arbeit umfassen im Einzelnen:

ein Übertragungsverfahren, das eine mehrstufige Modulation für überlagerte Funksignale ermöglicht und keine Frequenz- oder Phasensynchronisation benötigt,

ein Synchronisationsverfahren, das auf überlagerten Funksignalen beruht, und alle bestehenden Verfahren an Genauigkeit und Skalierbarkeit übertrifft,

ein Kanalzugriffsverfahren, das bisher nur auf drahtgebundenen Systemen möglich war; mit überlagerten Funksignalen kann es auch in drahtlosen Systemen eingesetzt werden,

ein verteiltes Datenfusionsprotokoll, das auf überlagerten Funksignalen basiert und dessen Prozesszeit unabhängig von der Anzahl der Teilnehmer ist; es kann somit als skalierungsfreies Protokoll verstanden werden,

umfassende Implementierungen und Anwendungen von *allen* inhaltlichen Beiträgen.

2 Analyse Zielplattform und Kommunikation

Die Kommunikation zwischen den Teilnehmern eines Funknetzes (hier mit Sensorknoten bezeichnet), kann unterschiedlichste Ausprägungen haben. Die wichtigsten Unterschiede bestehen zunächst in der Anzahl der Teilnehmer eines Kommunikationsvorgangs. Hier wird unterschieden, ob einer oder mehrere zu einem oder mehreren Teilnehmern kommunizieren. In dem Fall, dass mehrere Teilnehmer gleichzeitig zu einem Teilnehmer kommunizieren, kann es zu Überlagerungen von Informationen kommen. Will man diese Überlagerung bewusst erzeugen, so ist es daher notwendig, dass die Gleichzeitigkeit der Kommunikation sichergestellt wird. Dafür braucht man Zeitsynchronisationsverfahren. Im folgenden Kapitel werden bekannte Verfahren zur Zeitsynchronisation in drahtlosen Netzen analysiert. Weiterhin werden schon bekannte Verfahren zur Überlagerung von Funksignalen vorgestellt und ihre Anwendbarkeit auf die Zielplattform untersucht. Die Zielplattform stellt hier eine besondere Hürde dar, da sie extrem harte Anforderungen an die Komplexität und Realisierbarkeit stellt. Die Analyse der bekannten Verfahren von überlagerten Funksignalen und Verfahren zur Zeitsynchronisation zeigt, dass die herkömmlichen Ansätze für die Zielplattform nicht geeignet sind. Deshalb sind hier neue Verfahren nötig, welche die vorliegende Arbeit in späteren Kapiteln darstellen wird. Im folgenden Kapitel werden die Beiträge der vorliegenden Arbeit im internationalen, wissenschaftlichen Umfeld platziert und in Anforderungen und Zielen mit den bekannten Verfahren verglichen. Es wird auch der Systemaufbau vorgestellt, der durch die Analyse motiviert wird.

Für die Diskussion der Kommunikation zwischen drahtlosen Sensorknoten soll nun der Ablauf einer solchen Kommunikation genauer untersucht werden. Ein Sensorknoten, der eine Information übermitteln will, sendet diese auf der physikalischen Schicht aus. Die Prozesse in höheren Schichten wie Adressierung und Transport sind für diese Betrachtung nicht relevant. Deshalb soll hier auch nur die Kommunikation zwischen Knoten betrachtet werden, die direkte Funkverbindung haben, sich also in *Singlehop*-Distanz befinden. *Multihop*-Szenarien sollen ausgeschlossen werden, da sie als Zusammensetzung aus Singlehop-Prozessen dargestellt werden können. Für die Übertragung durch die Luft werden darüber hinaus keine Richtantennen in Betracht gezogen. Somit hat das Medium selbst eine verteilende (Rundruf-) Eigenschaft. Die Information, die ausgesandt wird, kann von all denjenigen Partnern empfangen werden, die im Sendebereich liegen. Die Natur der Funkübertragung verbietet eine paarweise Kommunikation. Zwar wird in herkömmlicher Netzwerktechnik oftmals Datenverkehr nur zwischen zwei ausgewählten Partnern betrieben. Dies ist jedoch eine Funktion einer höheren Schicht, so dass z.B. empfangene Datenpakete aufgrund von Zieladressen gefiltert werden. Prinzipiell wird oder kann jedoch jede Nachricht aus der Umgebung empfangen werden.

An einem Nachrichtenaustausch können auf beiden Seiten (Quelle und Ziel) ein oder mehrere Knoten beteiligt sein. Dies führt zu vier unterschiedlichen Kommunikationssituationen, die nun genauer beschrieben werden sollen. Besonderes Augenmerk liegt hier auf dem Entstehen von überlagerten Funksignalen, die Thema dieser Arbeit sind:

1 zu 1 Dies ist der Standardfall in der Literatur zu Mobilfunk. Ein Knoten verschickt ein Paket an einen Nachbarknoten. Häufig wird die Kommunikationstechnologie auf genau diesen Fall ausgelegt und optimiert. Die Eigenschaft des Mediums, Daten prinzipiell in einen ganzen Raumbereich hinein zu verschicken, wird dabei häufig ignoriert. Entwickelt man nun Verfahren, die die **1 zu 1**-Kommunikation unterstützen, so sind diese direkt auch für den nächsten Fall der **1 zu N**-Kommunikation anwendbar.

1 zu N Wie oben erwähnt, kann ein ausgesandtes Paket von allen denjenigen empfangen werden, die im Sendebereich liegen. Es ist die Verallgemeinerung des ersten Falls.

N zu 1 Wenn mehrere Sender eine Nachricht an ein Ziel verschicken, kann es zu Überlagerungen von Signalen kommen. Dies passiert dann, wenn die Signale der N Sender zur gleichen Zeit verschickt werden. Im Allgemeinen

kann dies absichtlich passieren (so wie es im weiteren Verlauf dieser Arbeit hauptsächlich der Fall ist) oder zufällig passieren. Je nach Modulation und Einsatz physikalischer Ressourcen kann dies auch zu Kollisionen führen.

N zu N Auch für dieses Szenario können überlagerte Signale relevant sein; es ist wiederum als Verallgemeinerung des Falls **N zu 1** zu verstehen.

Für die weiteren Betrachtungen wollen wir uns zunächst auf das **N zu 1**-Szenario beschränken. Es liegt also der Fall vor, dass mehrere Sender gleichzeitig Signale verschicken, die auf dem Medium zu Überlagerungen führen. Ein einzelner Empfänger soll nun – unter Verwendung nur seiner eigenen Ressourcen – aus dieser Überlagerung einen Vorteil ziehen. Wie dieser Vorteil aussieht, kann völlig unterschiedlich sein. Im Laufe dieser Arbeit werden mehrere, verschiedene Nutzungsmöglichkeiten solcher Überlagerungen vorgestellt werden.

Mit Hilfe von Abbildung 2.1 werden einige Arten von überlagerten Funksignalen unterschieden. Die Platzierung der vorliegenden Arbeit ist ebenfalls dargestellt und soll nun detailliert erläutert werden. Aus der Literatur sind besonders die *orthogonalen* Überlagerungstechniken bekannt. Orthogonalität soll in der Analyse als erstes Unterscheidungsmerkmal herangezogen werden. Orthogonalität wird oft in verschiedener Ausprägung verstanden. Orthogonalität kann zwischen Frequenz, Zeit oder einer Kombination beider entstehen (wie z.B. CDMA-Systeme). Bei Systemen, die Funksignale orthogonal übertragen, handelt es sich um die klassischen Übertragungssysteme wie Mobilfunk oder auch drahtgebundene Rechnernetzwerke. Eine ausgesandte Information soll im Empfänger stets individuell detektiert und dekodiert werden können. Deshalb ist eine Orthogonalität der Informationen auf dem Medium erforderlich. Nur so können Informationen verschiedener Quellen im Empfänger wieder klar getrennt werden.

Bei nicht-orthogonaler Übertragung ist die Auflösung der überlagerten Signale und die Zuordnung auf die Quellen prinzipiell nicht vollständig möglich. Die Überlagerung dient hier nicht der gleichzeitigen Nutzung des Kanals, sondern anderen Anwendungen, die im Laufe dieser Arbeit vorgestellt werden. Wenn Funksignale nicht orthogonal überlagert werden und darüber hinaus keinerlei Koordination besteht, kommt es zu sogenannten Kollisionen. Das heißt, die ausgesandten Funksignale vermischen sich mit Signalen anderer Kommunikationspartner, das Gemisch ist nicht mehr trennbar und das Wellenchaos trägt kaum noch brauchbare Informationen. Die Datenübertragung bedarf dann einer vollständigen Wiederholung. Eine solche unkontrollierte Überlagerung kann z.B. in FDMA-Systemen

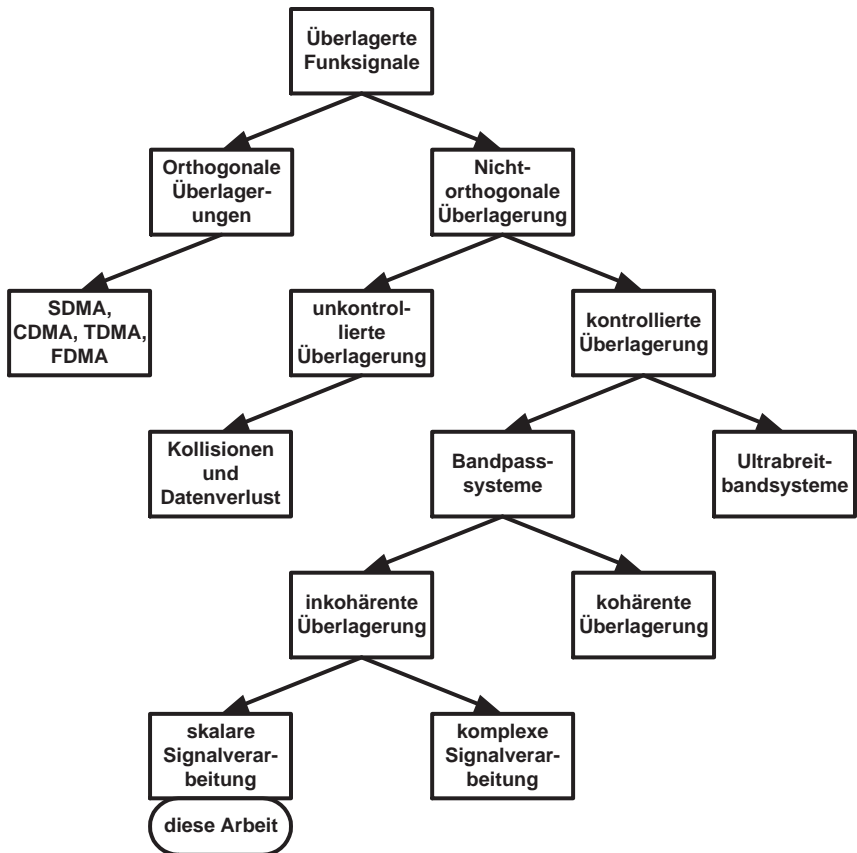


Abbildung 2.1. Überblick über verschiedene Arten von überlagerten Funksignalen

auftreten, wenn zwei Sender auf der gleichen Trägerfrequenz übertragen, weil die Absprache über die Kanäle fehlgeschlagen ist.

Unter kontrollierter, aber nicht-orthogonaler Überlagerung wird hier verstanden, dass ohne die Orthogonalität zwar die ausgesandten Signale nicht mehr vollständig voneinander trennbar sind, aber die Überlagerung absichtlich ausgeführt wird und einem bestimmten Zweck dient. Das heißt natürlich auch, dass es sehr wohl weitere Festlegungen über diese Signale geben kann, auch wenn diese nicht orthogonal sind. Dies kann die Impuls- oder Wellenform der Signale, ihre Energie, Dauer, Startzeiten etc. betreffen. Nur bei einem gewissen Mindestmaß an Kontrolle können nicht-orthogonale Signale brauchbare Ergebnisse erzielen.

Als nächstes Unterscheidungskriterium wurde der Bandbreitenbedarf gewählt. Ultrabreitbandsignale¹ sind geeignet, um Überlagerungen von Signalen zu erzeugen. Da die Empfänger von Ultrabreitbandsignalen typischerweise auf Integration des Zeitsignals basieren, sind sie die ideale Vorbereitung für überlagerte Signale, die ja auch aus der Signalsumme, also der Integration, Vorteil ziehen. Systeme basierend auf Ultrabreitbandsignalen sind jedoch wenig verbreitet und deren Lizenzierung ist weltweit noch kaum geregelt. Dies liegt nicht zuletzt an der schwierigen Koexistenz von mehreren Ultrabreitbandsystemen. Die Orthogonalität verschiedener, gleichzeitig betriebener Systeme (wie z.B. die Trennung durch Funkkanäle) ist nicht ohne weiteres möglich.

Ein Extremfall von kontrollierten, überlagerten Signalen ist die kohärente Überlagerung. Sie bedeutet, dass Signale, die überlagert werden, in Phase, Frequenz und Symbol-Startzeit so exakt übereinstimmen, dass es zu stabiler, konstruktiver Interferenz kommt. Kohärente Überlagerungen (Abbildung 2.2) führen zu konstruktiver Interferenz und können so die Signalstärke erhöhen. Dafür ist es allerdings nötig, Phase, Frequenz und Startzeit von Signalen verschiedener Quellen genau aufeinander abzustimmen. Ist dies nicht der Fall, so kommt es zu zeitlich und örtlich wechselnden Verstärkungen und Auslöschungen, sogenannten „Schwebungen“ (siehe Abbildung 2.3) und „Funklöchern“. Ein kohärent arbeitendes System kann daher auch stets nur zu einem gegebenen Empfangsort Kohärenz sicherstellen, da die Laufzeitunterschiede der Signale aus verschiedenen Quellen die Lage der Wellenzüge in der Zeit — und damit die Kohärenz — beeinflussen. Zusätzlich zu Laufzeitunterschieden kann es zu weiteren Effekten, wie Mehrwegeausbreitung und Dämpfungen kommen. Dies wird allgemein in

¹Unter Ultrabreitbandsignalen versteht man (laut FCC und ITU) Signale, die entweder mehr als 500Mhz Bandbreite besetzen oder deren Bandbreite mehr als 20% ihrer Mittenfrequenz beträgt

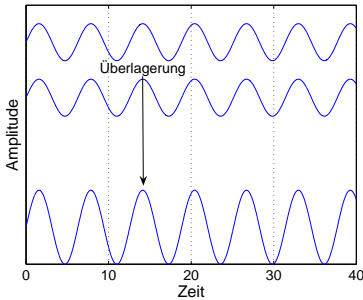


Abbildung 2.2. Kohärente Überlagerung führt zu konstruktiver Interferenz

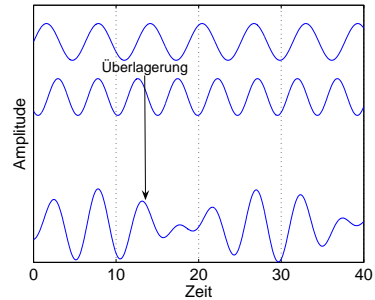


Abbildung 2.3. Inkohärente Überlagerung führt zu Schwebungen

Kanalmodellen zusammengefasst. Inkohärente Überlagerungen müssen deshalb mit den Schwierigkeiten all dieser Einflüsse des Kanals zurechtkommen. Daher müssen Signalformung und Detektion die Effekte des Kanals berücksichtigen. Dies gilt natürlich allgemein für jedes Datenübertragungssystem. Für Systeme, die überlagerte Funksignale einsetzen, ist es zusätzlich nötig, auch die Effekte der Summation der Signale auf dem Kanal zu berücksichtigen.

Nun soll zunächst die Zielplattform und der Anwendungsbereich dieser Arbeit näher spezifiziert werden. Der Anwendungsbereich gibt deutliche Einschränkungen vor und legt somit auch die Klasse von überlagerten Funksignalen fest, die im weiteren Verlauf diskutiert werden wird.

2.1 Die Zielplattform

Überlagerte Funksignale können für Funkssysteme jeder Ausprägung verschiedene Vorteile bringen. Im Rahmen dieser Arbeit sollen jedoch nur ganz spezielle Funkssysteme mit sehr wenig Ressourcen betrachtet werden. Es handelt sich also um kostengünstige und einfache Endgeräte. Diese Klasse von Geräten wird nun als die *Zielplattform* der vorliegenden Arbeit betrachtet. Alle Mechanismen und Teilbeiträge der Arbeit sind auf diese Zielplattform ausgerichtet. Sensornetzwerke, oder genauer gesagt: *Sensorknoten* gehören zur Zielplattform und sind der technische Hauptanwendungsbereich der vorliegenden Arbeit. Die Zielplattform

umfasst jedoch auch noch einfachere Geräte wie RFID oder sogar Polymerelektronik als Extrembeispiele. Die Zielplattform soll nun zunächst auf technischer Seite charakterisiert werden:

Speicher. Die Zielplattform verfügt über extrem wenig Speicher². Es ist sogar denkbar, dass die Zielhardware überhaupt keinen Arbeitsspeicher enthält, so wie viele passive RFID Systeme und alle Prozesse durch feste Zustandsabläufe festgelegt sind.

Taktrate. Die Taktraten der Prozessoren oder Automaten sind typischerweise nur einige kHz bis wenige MHz. Der Hauptgrund ist der Energieverbrauch. Um diesen niedrig zu halten, muss die Taktrate auch so niedrig wie möglich gehalten werden.

Prozessorleistung. Die Prozessorleistung ist ebenfalls gering. Dies impliziert, dass keine Co-Prozessoren vorhanden sind, keine Fließkommaarithmetik oder ähnlich unterstützende Module vorhanden sind. Digitale Signalverarbeitung wird von Hardwareseite nicht unterstützt.

Energieversorgung. Die Knoten sind mobil und so ist eine kabelgebundene Energieversorgung nicht möglich. Der Einsatz von Batterien ist deshalb üblich. Um die Lebensdauer hoch zu halten, ist Energieverbrauchsoptimierung in allen Systemteilen erforderlich und ein wichtiger Gestaltungsspekt.

Diese Einschränkungen der technischen Leistungsfähigkeit führen zu einer stark reduzierten Gesamtfunktionalität. Die geringe Taktrate des Empfängers mit geringem Speicher und schwacher CPU führt dazu, dass sich die Zielplattform als ein sehr leistungsschwaches System darstellt. Aus den technischen Einschränkungen können nun die Einschränkungen der Funktionalität abgeleitet werden:

Keine komplexe (De-) Modulation. Es wird angenommen, dass ein solches Verhalten wegen des sehr einfachen Aufbaus der Hardware nicht möglich ist. Dies ist ein schwerer Einschnitt in die Möglichkeiten des Übertragungssystems. Der Großteil der vorhandenen Theorie für Signalformung und Detektion kann nicht angewendet werden. Phasenmodulationen sind vollständig ausgeschlossen. Es bleibt also nur noch die Amplitude oder Frequenz als Modulationsparameter übrig. Für die Demodulation sind dann sehr einfache Verfahren möglich wie frequenzselektive Filter (evtl. Analog) oder

²typischerweise einige hundert Byte RAM und wenige kByte ROM

einfache Amplitudendenmodulation über die Einhüllende (ebenfalls durch einfache analoge Schaltung realisierbar).

Keine hohe Abtastrate. Die leistungsschwache Hardware kann keine hohen Abtastraten der Basisbandsignale leisten. In den meisten Fällen (allen existierenden Plattformen für Sensorknoten) ist die Geschwindigkeit der möglichen A/D Wandlung nicht einmal ausreichend, um das Abtasttheorem für die empfangenen Funksignale zu erfüllen. Somit ist eine Softwarelösung zur Verarbeitung der Funksignale im Basisband nicht darstellbar.

Möglichst wenig digitale Signalverarbeitung. Aufgrund der geringen Prozessorleistung sind aufwendige Filterketten und ähnliche digitale Signalverarbeitung nicht möglich. Die bereits angesprochenen Probleme mit der A/D Wandlung verbieten hier ebenfalls eine Lösung in Software.

Bandpasssignale. Die Verwendung von Bandpasssignalen auf dem Funkkanal ist eine Anforderung, die sich automatisch aus heutiger Lizenzlage ergibt. Das Funkmedium ist nur in bestimmten Kanälen für z.B. Sensornetzwerke freigegeben. Somit ist der Arbeitsbereich auf diese freien Funkbänder eingeschränkt. In Deutschland sind dies typischerweise Funkbereiche um 433MHz, 868MHz und 2.4GHz.

Diese funktionalen Einschränkungen beschreiben ein Übertragungs-Frontend, das sehr leistungsschwach ist. Eine mögliche technische Darstellung eines solchen Systems wäre ein Bandpasssystem, das z.B. auf 433MHz arbeitet und mit einem OOK oder 2-FSK Übertragungsverfahren arbeitet. Es treten keine komplexwertigen Signalpfade in der Verarbeitung auf und die Empfangsfilter sind analog aufgebaut. Beispiele für Sensornetzwerke, die aus solchen Komponenten aufgebaut sind, sind z.B. pPart oder cPart Particle Computer³ oder auch die Berkeley Motes⁴.

Die genannten technischen Eigenschaften und Charakteristika können für die Zielplattform auch als *Anforderungen* verstanden werden: Endgeräte, die mindestens die beschriebene technische Leistungsfähigkeit erbringen können, zählen zur Zielplattform dieser Arbeit und die hier vorgestellten technischen Beiträge sind geeignet, um auf diesen Endgeräten implementiert zu werden.

³<http://www.particle-computer.de>

⁴<http://www.crossbow.com>

2.2 Bekannte Verfahren für überlagerte Funksignale

Überlagerte Funksignale werden in der Literatur bereits seit einiger Zeit diskutiert. Im Prinzip könnte man sogar Arbeiten aus den Anfängen der Informationstheorie in den 40er Jahren dazu zählen. In [52] entwickelten die Autoren Methoden, um überlagerte Codes zu detektieren. Diese Idee lässt sich auch sehr weit auf frühe Arbeiten von Kautz und Singleton ([39]) zurückdatieren. Sie schreiben dort: „A binary superimposed code consists of a set of code words whose digit-by-digit Boolean sums ($1 + 1 = 1$)“. Diese Idee wird im Kapitel 7 unter dem Gesichtspunkt überlagerter Signale weiter betrachtet werden. Weitere, fundamentale Arbeiten zu überlagerten, analogen Signalen für multi-user binary-adder-Kanäle können in [10] gefunden werden. Die Autoren diskutieren physikalische Einschränkungen von überlagerten Signalen, die nicht orthogonal sind.

In der neueren Literatur tauchen überlagerte Funksignale als eine Thematik aus der Forschung über *cooperative transmission* auf. Diese Bezeichnung wird allerdings in verschiedenen, unterschiedlichen Kontexten eingesetzt. Unter *cooperative transmission* wird meistens die Grundidee verstanden, dass Relaisstationen gesendete Pakete weiterleiten, um die Detektion im Empfänger zu verbessern. In solchen Systemen wird vor allem der *diversity gain* ausgenutzt um durch verschiedene, unabhängige Signalpfade zum Empfänger die Empfangsqualität im Empfänger zu verbessern. Aus Systemsicht gibt es unterschiedliche Arten, wie die Informationen weitergeleitet werden können:

Amplify and forward: Hier wird das bei der Relaisstation eingetroffene Paket verstärkt und wieder ausgesandt. Das Paket wird lediglich als Signal verstanden. Es findet keine Detektion oder Demodulation statt – nur eine Verstärkung.

Decode and forward: Hier dekodiert die Relaisstation die eintreffende Information. Nur wenn das Paket fehlerfrei empfangen wurde, werden die Signale auf der physikalischen Schicht identisch neu aufgebaut und das Paket wieder ins Medium versandt.

Cooperative coding: Relaisstationen dekodieren die eintreffenden Pakete, verwenden aber verschiedene Codes und Redundanzen bei der Neugenerierung, um die Fehlererkennung und Fehlerkorrektur im Empfänger zu verbessern. Die verschiedenen Signalpfade zwischen Quelle und Senke werden also dafür genutzt, eine größere Redundanz zu erreichen.

In [48] werden diese drei Arten von kooperativen Verbindungen diskutiert und Kodierungsstrukturen hergeleitet und in [62] werden die entsprechenden Fehleraten verglichen. In [66] geben die Autoren einen allgemeinen Ausdruck für die Symbolfehler in *amplify and forward* Systemen an, der mithilfe einer eigens dafür entwickelten Berechnungsmethode dargestellt wird. Weiterhin wird in [15] ein genereller Maximum-Likelihood-Empfänger für binäre FSK in einem *decode and forward*-System vorgestellt.

Diese Arbeit versteht den Begriff *cooperative transmission* in einer leicht anderen Art. Die Idee ist, dass Sensorknoten gleichzeitig Signale versenden, um mit dieser *kooperativen Überlagerung* einen zusätzlichen Nutzen zu erzeugen. Andere wissenschaftliche Arbeiten, die sich auch mit genau diesem Aspekt befassen, sind in zwei Forschungsrichtungen aufgeteilt. Diese Forschungsrichtungen lassen sich durch die Systemvoraussetzungen unterscheiden:

Coherent cooperative transmission: Bei kohärenter Überlagerung von Signalen wird angenommen, dass alle Sender mit dem Empfänger Kohärenz aufweisen. Dies bedeutet, dass Träger- und Zeitsynchronisation ideal erreicht sind. In [78] ist eine Systembeschreibung für eine solche Überlagerung dargestellt. Die Autoren schlagen einen Algorithmus vor, wie diese komplizierte Synchronisation in einem verteilten System tatsächlich hergestellt werden kann. Die Gewinne für die Detektion sind dann erheblich.

Non-coherent cooperative transmission: Funkknoten überlagern ihre Signale inkohärent. Phasen φ_l und Träger ω_l sind nicht ideal synchronisiert. Dies ist der Forschungsbereich der vorliegenden Arbeit.

2.2.1 Opportunistic Large Arrays

Die CRISP (Communications Research in Signal Processing) Gruppe der Cornell University, Ithaca, New York, USA, hat in den letzten Jahren unter der Leitung von Prof. Anna Scaglione einige Arbeiten hervorgebracht, die in Ausrichtung und Zielsetzung der vorliegenden Arbeit sehr ähnlich sind. Diese und die Arbeiten der CRISP Gruppe wurden unabhängig voneinander entwickelt und zeitgleich veröffentlicht. Deshalb sollen diese Arbeiten jetzt detaillierter vorgestellt werden, um die Unterschiede herauszuarbeiten und die verschiedenen Beiträge voneinander zu separieren. Der Hauptunterschied zwischen der vorliegenden Arbeit und den Arbeiten aus der CRISP Gruppe sind die Systemvoraussetzungen und Praxisrelevanz. Die Arbeiten der CRISP Gruppe sind sehr informationstheoretisch

ausgelegt und verwenden in den Systemvoraussetzungen allgemeine Annahmen wie komplexe Datenverarbeitung und Einsatz von breitbandigen Signalen. Diese Arbeit befasst sich hingegen besonders mit Mechanismen und Verfahren, die kooperatives Überlagern für sehr einfach aufgebaute Hardware ermöglicht. Auch werden im Gegensatz zu den Arbeiten der CRISP Gruppe alle vorgestellten Verfahren einer kritischen Prüfung in Implementierungen auf Hardware unterzogen. Die Arbeiten zu kooperativer Überlagerung umfassen in der CRISP Gruppe auch

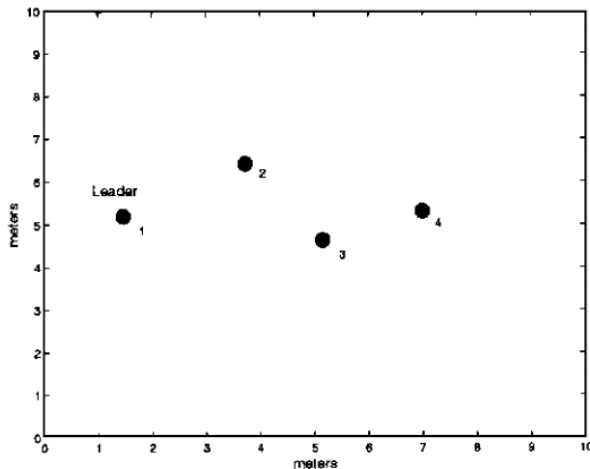


Fig. 2. Simple example of OLA: Four-node case.

Abbildung 2.4. Die Position der Knoten im Beispiel der Opportunistic Large Arrays, aus [71, S. 2085]

Anwendungen wie Transportmechanismen (siehe z.B. [75]). Ein Teil der Veröffentlichungen verfolgt eine zentrale Idee: die der *opportunistic large arrays*. Die Idee ist dort, ein Signal, das als breitbandiger „Impuls“ vorliegt, auszuschicken. Stationen, die diesen Impuls empfangen, versenden instantan auch einen Impuls, so dass es zu einer Überlagerung kommt und sich die Sendeleistungen addieren. Abbildungen 2.4 und 2.5 zeigt, wie der Ablauf des OLA-Systems funktioniert: In diesem Szenario beginnt Node 1 mit dem Versenden eines Impulses. Bei Node

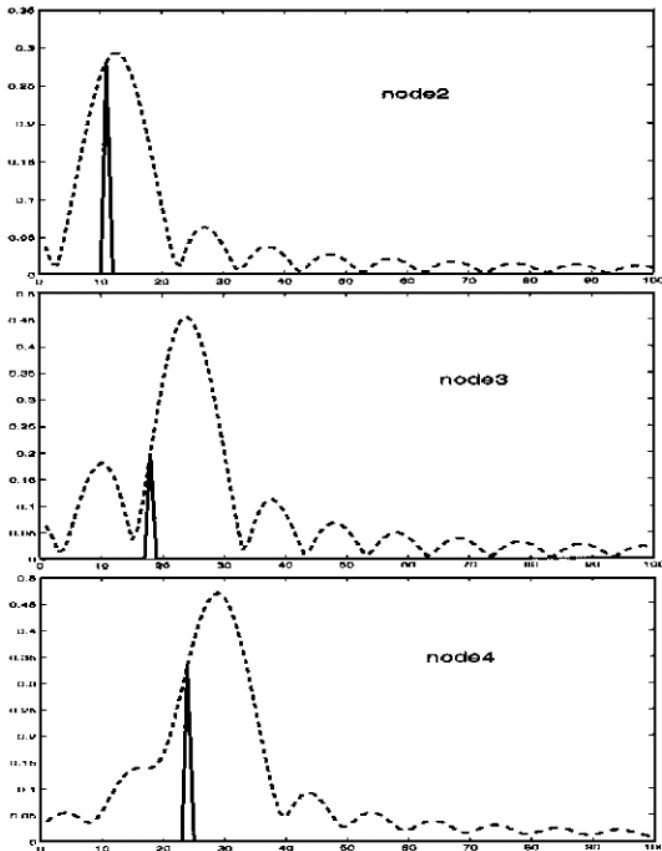


Fig. 3. Illustration of the received signal at nodes 2-4, corresponding to Fig. 2. The spikes represent the point where the SNR threshold is met and, thus, the beginning of transmission at that node, i.e., the firing instant.

Abbildung 2.5. Der zeitliche Verlauf der Signalisierung für Opportunistic Large Arrays, aus [71, S. 2085]

2 wird dies empfangen, und als das SNR eine gewisse Schwelle überschreitet, sendet Node 2 auch einen Impuls aus. Das ist mit der Spitze in den unteren Darstellungen von Abbildung 2.5 angedeutet. Etwas später sind auch für Node 3 und 4 die SNR-Schwellen erreicht, und das Signal breitet sich wie eine Welle aus⁵.

Im Unterschied zur vorliegenden Arbeit werden bei Signalgenerierung und Detektion zunächst keine Einschränkungen vorgegeben. Ähnlich zu 3.6 wird nun das empfangene Signal wie folgt formuliert [71, Formel (2)]:

$$s_{i,m}(t) = \sum_{n=1}^N A_{i,n}(t) p_m(t - \tau_{i,n}(t)), \quad m = 0, \dots, M-1 \quad (2.1)$$

p_m ist die Impulsform, $A_{i,n}$ beinhaltet Kanal, Sendeleitung und Dämpfung und $\tau_{i,n}$ reflektiert die unterschiedlichen Anfangszeiten durch den Welleneffekt. Die weiteren Herleitungen setzen nun komplexe Modulationen und Demodulationen voraus, die für diese Arbeit nicht relevant sein sollen.

Eine weitere Arbeit ([74]) der CRISP Gruppe, die die Detektion von überlagerten Signalen betrifft, ist für diese Arbeit interessant. Dort wird mit einem Modell, das im Wesentlichen (2.1) entspricht, ein optimaler Detektor hergeleitet. Wie auch in dieser Arbeit (siehe Kapitel 3.1), wird der zentrale Grenzwertsatz benutzt, um die Statistik im Empfänger herzuleiten. Mit einem generalisierten Likelihood-Test wird ein Empfänger hergeleitet, der auf den komplexen Abtastwerten eines „sliding window“ über eine gegebene Länge angewandt wird. Diese Herangehensweise führt zu einer hohen Detektionsleistung. Unter den Voraussetzungen von Kapitel 2.1 ist jedoch ein solches System nicht realisierbar. Es scheitert gleich an mehreren Stellen. Die Signalverarbeitung mit einem „sliding window“ ist sehr aufwendig und benötigt enorme Rechenleistung. Darüber hinaus ist auch eine Generierung von Signalen im Komplexen in der vorliegenden Arbeit nicht vorgesehen. Des Weiteren sind auch Training und Kanalschätzung nicht vorgesehen oder beschränken sich auf einfache Verfahren, die z.B. lediglich den Betrag der Dämpfung in Betracht ziehen.

2.2.2 Spectrum Pooling

Eine andere Arbeit [24] von J.Hillenbrand, T. Weiss und F. Jondral, die im Zusammenhang mit dem deutschen BMBF-HyEff Forschungsprojekt an der Universität Karlsruhe im Jahr 2002 entstanden ist, befasst sich ebenfalls mit überlagerten

⁵wie „la ola“ in Sportstadien, daher auch der Name

Funksignalen. Der Fokus dieser Arbeit bezieht sich auf ein Szenario, bei dem eine Funkkanalbelegung erkannt werden soll. Es handelt sich also nicht um eine *Datenübertragung* mit Hilfe von überlagerten Funksignalen, sondern lediglich um eine Signaldetektion. Dies geschieht, indem die Leistung auf diesem Kanal gemessen wird. Um die gesamte Erkennung zu verbessern, geht man hier den Weg der verteilten Erkennung: Alle anwesenden Funkteilnehmer führen eine Erkennung von Zeit zu Zeit durch. Falls einer der Teilnehmer erkannt hat, dass auf dem zu betrachtenden Kanal ein Signal ist, dann sendet auch er zusätzlich ein zufälliges Signal in diesen Kanal, um die Empfangsleistung für die anderen Teilnehmer und so ihre Detektionswahrscheinlichkeit zu erhöhen. Diese wurde als „boostig-protocol“ in [83] vorgestellt. Abbildung 2.6 zeigt das Anwendungsszenario. Die

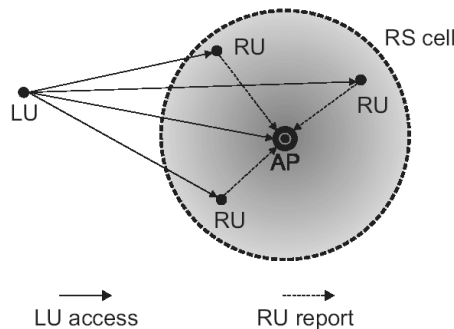


Abbildung 2.6. Verstärkung von Signalen für die verteilte Detektion; aus [83]

Signale, die mit „RU report“ bezeichnet sind, sind diejenigen Signale, die von Teilnehmern kommen, die ein Funksignal entdeckt haben und dieses jetzt verstärken, um anderen die Detektion zu erleichtern. Da die Quelle des ursprünglichen Signals (LU) nicht in informeller Funkverbindung mit den anderen Teilnehmern (RU) steht, ist auch keine Synchronisation in Träger, Phase oder Zeit möglich. In der Arbeit [24] handelt es sich um die Herleitung einer Detektionsvorschrift für ein Signal, dass eine unbekannte Quelle aussendet. Die Detektion beruht auf einem Likelihood-Test und die Entdeckungs- und Fehlarmswahrscheinlichkeiten werden nach dem Neaman-Pearson-Kriterium berechnet.

Der Bezug zu der vorliegenden Arbeit besteht darin, dass sich die Autoren ebenfalls mit der Detektion von überlagerten Funksignalen befassen. Sie setzen allerdings voraus, dass das Empfangssignal im Basisband abgetastet und komplex verarbeitet wird. Komplexe Signalverarbeitung ist im Rahmen der vorliegenden Arbeit nicht vorgesehen. Auch befassen sich die Arbeiten über Spectrum Pooling nur mit der Detektion und nicht mit der Signalformung für die Überlagerung.

2.2.3 dFSK

„dFSK“ steht für „distributed frequency shift keying“ und wurde von An-Swol Hu und Sergio Servetto 2003 veröffentlicht [31]. Sie beschreiben ein Modulationsverfahren, das auf überlagerten Funksignalen basiert. Durch die Ähnlichkeit zu normaler FSK-Modulation ergab sich die Namensgebung: „The aggregate waveform that results from superimposing a large number of these weak transmissions has a pre-specified set of zero-crossings“ [31]. Das Anwendungsszenario ist das bekannte Sensor-Reach-Back-Problem. Die Überlagerung basiert auf einer hochsynchronen Signalüberlagerung, die mit Hilfe einer speziellen Synchronisation erreicht wird. Diese ist ebenfalls von den Autoren in [32] vorgestellt. Die Synchronisation erzeugt eine phasengenaue Synchronisation von ausgesendeten Pulsen. Die darunterliegende Theorie baut auf verteilten, gekoppelten Oszillatoren auf, die über eine verteilte Regelung in Synchronisation gebracht werden. In Abbildung 2.7 ist die Wellenform nach der Überlagerung abgebildet. In der Arbeit werden keine Kanalmodelle verwendet und es wird nicht speziell auf die Anforderungen der Wellenformgenerierung eingegangen. Zwar ist das System als Bandpasssystem implementierbar, dennoch erfordert es eine exakte Herstellung von Wellenformen im Transmitter. Dafür sind sehr exakte Synthesebausteine erforderlich. Solch aufwendige Verfahren sind auf der Zielplattform der vorliegenden Arbeit nicht möglich.

Weiterhin stellen sich zwei Fragen, die für das vorgestellte Verfahren nicht gelöst werden: Der Aufwand, phasengenaue Signale auf der Trägerfrequenz eines Systems zu erzeugen, ist erheblich, und der Einfluss des Kanals wird nicht diskutiert. Der Träger selbst müsste aus einem Wellenformgenerator wie z.B. einer IFFT erzeugt werden. Dies ist z.B. in OFDM-Verfahren bekannt. Der Transceiveraufbau analog zu OFDM-Systemen ist jedoch hier nicht möglich, da dort das Signal im Basisband (also bei niedrigen Frequenzen) per IFFT erzeugt und dann auf den Träger aufmoduliert wird. Der Empfänger arbeitet umgekehrt. Dies ist für das hier diskutierte Synchronisationsverfahren und dFSK nicht möglich. Die

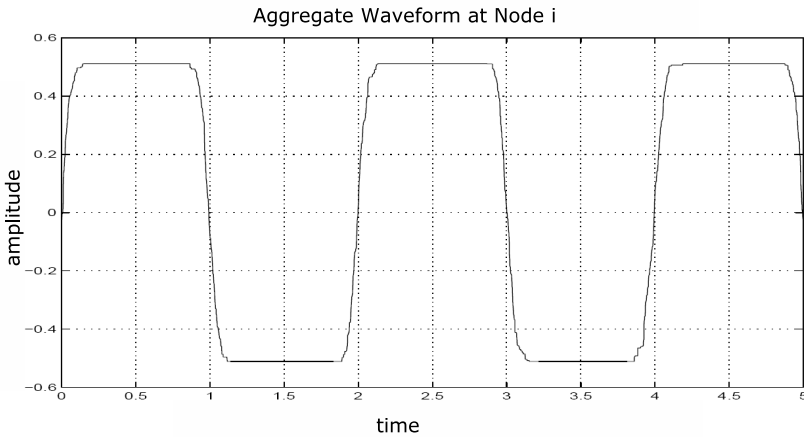


Abbildung 2.7. dFSK: Ergebnis der Überlagerung; aus [31, Abbildung 2]

Verfahren basieren ja auf Überlagerung, und diese findet im Bandpass statt und nicht im Basisband. Deshalb muss auch die Nachregelung für die Synchronisation auf der tatsächlich ausgesandten Wellenform stattfinden und der Einfluss der Modulation auf die Trägerfrequenz beachtet werden. Die Autoren schlagen deshalb im Ausblick in [32] vor, statt exakt pulsgeformten Signalen Ultrabreitbandsignale einzusetzen. Es wird jedoch nicht weiter diskutiert, ob das System mit seinen Eigenschaften direkt auf Ultrabreitbandtechnologie verwendbar ist. Sowohl die aufwendige Erzeugung der pulsgeformten Signale als auch der Einsatz von Ultrabreitbandsignalen ist daher deutlich theoretischer Natur. Die Schwierigkeiten der Implementierung und der Einfluss von Kanalmodellen werden nicht diskutiert.

2.2.4 Kohärente Überlagerung

Yung-Szu Tu und G. Pottie befassen sich in ihrer Arbeit [78] damit, wie Signale, die von verteilten Sensorknoten ausgesandt werden, kohärent überlagert werden können. Sie verwenden zur Erklärung die Abbildung 2.8. Diese Abbildung entspricht sehr stark den Diskussionen, die auch in der vorliegenden Arbeit angestellt werden. Die Ähnlichkeit zu Abbildungen 2.2 und 2.3 ist deutlich. Die Autoren

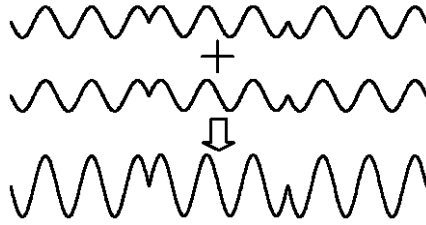


Abbildung 2.8. Kohärent überlagerte Signale; aus [78, Abbildung 1, Seite 2]

befassen sich jedoch ausschließlich mit der kohärenten Überlagerung. Sie stellen unter anderem fest, dass dadurch ein deutlich höherer Leistungsgewinn erzeugt werden kann. Angenommen, die Dämpfungen α und Sendeleistungen P_{tx} zum Zielpunkt wären gleich, so wäre die gesammelte Leistung einfach αNP_{tx} . Wenn aber die Signale kohärent überlagert werden, so ist der Gewinn *am Zielort* deutlich höher:

$$P_{av} = (N\sqrt{\alpha P_{tx}})^2 = \alpha N^2 P_{tx} \quad [78, \text{Formel}(2)] \quad (2.2)$$

Die Motivation wird aus einem Szenario eines Mars-Orbiter erbracht. Der Orbiter kann mit Hilfe von Sonnensegeln genug Energie sammeln, um die abgesetzten Mars-Erkunder anzufunkeln. Die mobilen Erkunderroboter können ihrerseits alleine die hohe Sendeleistung nicht aufbringen, um an den kreisenden Orbiter zurückzusenden. Also kombinieren sie ihre Leistungen. Für die kohärente Überlagerung umfasst die Synchronisation Träger, Phase und Symbolzeiten. Dafür gibt es individuelle closed-loop Regelkreise zwischen dem Master (Orbiter) und den Slaves (den Erkunderrobotern). Um dies zu erreichen, werden zwischen Slaves und Master Trägersignale ausgetauscht, die jeweils Vielfache voneinander sind und aus dem gleichen Oszillator (durch Teilung und Vervielfachung) erzeugt werden. Zusätzlich wird das System noch detaillierter für CDMA-Systeme untersucht und die nötigen Regelkreise werden vorgeschlagen. Die verwendeten Kanalmodelle sind AWGN⁶ und Flat-Fading. Das vorgeschlagene Synchronisationsverfahren setzt einen direkten Verbindungskanal zwischen Master und Slaves voraus. Somit findet dieses Verfahren sicherlich im Motivationsbeispiel derselben Arbeit [78] nicht direkt Anwendung. Es ist vielmehr ein anderer Punkt interessant, der

⁶AWGN steht für *additional white gaussian noise* und ist ein Kanalmodell, bei dem dem Nutzsignal eine additive Rauschkomponente mit gaußscher Statistik zugeführt wird.

auch in dieser Veröffentlichung angesprochen wird. Mit höherer Empfangsleistung — die durch die Zusammenschaltung mehrerer Slaves entsteht — steigt der Rauschabstand (SNR) im Empfänger und damit erhöht sich auch die Kanalkapazität zwischen Slaves und Master.

2.2.5 Kohärente Überlagerung zweier mobiler Quellen

In [36] haben D.R. Brown III, G.B. Prince und J.A. McNeill ein Synchronisationsverfahren vorgeschlagen, das speziell für mobile Szenarien geeignet ist. Sie erreichen Synchronisationszeiten von unter 1000 Trägerzyklen (also z.B. 800MHz: $1.25\mu s$). Die verwendeten Kanalmodelle reichen über einfache Einwegeausbreitung (Single-Path-Fading) bis zu zeitvarianten Modellen und Mehrwegeausbreitung. Hier wird ein Rice-Kanalmodell verwendet. Das Synchronisationsverfahren degradiert allerdings deutlich, wenn der Parameter K des Rice-Kanals gegen 1 geht. Dieser Parameter gibt das Verhältnis zwischen direktem Signal und Mehrwegesignal im Empfänger an. Zusammengefasst heißt dies, dass das Verfahren für Mehrwegeausbreitung eher nicht geeignet ist. Mehrwegeausbreitung ist jedoch ein wichtiger Aspekt in der vorliegenden Arbeit. Das System von Brown et al. ist aus zwei Sendern aufgebaut, die sich auf ein gemeinsames Ziel synchronisieren. Realisiert werden soll das System mit je zwei Frequenzsynthese-PLLs in den Sendern und bidirektionalen Funkverbindungen zwischen allen Teilnehmern. Größere Systeme mit mehr Teilnehmern können so nicht ohne weiteres realisiert werden, was aber in der vorliegenden Arbeit höchste Relevanz hat, da die Kommunikationsszenarien von Sensornetzwerken stets eine große Anzahl von Funkteilnehmern umfassen. Im Vergleich zur vorliegenden Arbeit ist der Systemaufbau auch deutlich komplexer. Synthesebausteine, optimale, komplexe Parameterschätzer etc. sind nötig.

2.2.6 Beam-forming mit Hilfe von kohärent überlagerten Funksignalen

In der Arbeit [11] von G. Barriac, R. Mudumbai und U. Madhow geht es darum, Signale von verschiedenen Sensorknoten phasengenau zu überlagern, um damit eine gerichtete Antenne zu imitieren. Der interessante Aspekt hierbei ist weniger die tatsächliche technische Ausführung und Lösung, als der Gedanke verteilter, synchronisierter Antennen an sich. Besonders für Systeme in niedrigen Frequenzbereichen unterhalb 100MHz ist die Wellenlänge im Meterbereich. Das

bedeutet, dass auch die Dimensionen für Richtantennen sich in der Größenordnung von Metern bewegen. Für Sensornetzwerke, die mit Kleinstrechner und Batterien arbeiten, ist es aus der Sicht des Systemdesigns und der Anwendungen völlig unpassend, ein Gerät, dass selbst nur wenige Zentimeter groß ist, mit einer metergroßen Antenne zu versehen. Durch den Einsatz verteilter Sensorknoten mit jeweils einer kleinen Antenne können aber über die Synchronisation Antennenarrays aufgebaut werden, um eben eine bestimmte Richtcharakteristik zu erzeugen. In Abbildung 2.9 ist im unteren Bereich zu sehen, wie sich die Antennenkeule

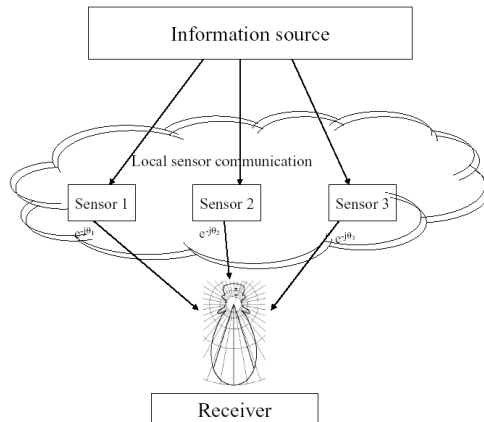


Abbildung 2.9. Kohärente Überlagerung von Signalen zur Erzeugung einer Richtcharakteristik; aus [36, Abbildung 8, Seite 8]

ausprägt. Damit die Überlagerung für beam-forming genutzt werden kann, muss auch bei diesem System exakte Synchronität von Trägern und Phasen erreicht werden. Die Autoren lösen dieses Problem nicht, sondern schlagen vor, aus der bekannten Distanz zum Ziel und der Ausbreitungsgeschwindigkeit die Laufzeiten zu berechnen und somit die Phasendifferenz zu schätzen. Die Frage der Genauigkeit der Phasenschätzung wird somit auf die Frage der Genauigkeit der Ortschätzung der Sensorknoten zurückgeführt. Für den dort skizzierten Anwendungsfall sind Ortsgenauigkeiten im 10cm-Bereich nötig, die sicherlich mit heute verfügbaren Lokationstechniken (z. B. Ultraschall) darstellbar sind. Signallaufzeiten in den Geräten werden zunächst nicht betrachtet. In Simulationen zeigen die Auto-

ren, dass das System prinzipiell funktioniert, setzen allerdings voraus, dass die Kanalkoeffizienten zwischen den Quellen und dem Ziel alle betragsmäßig identisch sind.

Die exakte Steuerung der Phasen und Frequenzen der ausgesendeten Signale erfordert wieder eine aufwendige Signalverarbeitung und -synthese. Dies ist auf der Zielplattform der vorliegenden Arbeit nicht möglich.

2.2.7 Virtuelle Radarbilder mit überlagerten Funksignalen

Die letzte, hier vorgestellte Arbeit [8] stammt von den Autoren Bharath Anathasubramaniam und Upamanyo Madhow. Ziel dieser Arbeit ist es, ein Radarbild eines physikalischen Phänomens zu erzeugen. Allerdings werden hierfür weder ein Radarstrahl noch reflektierende Objekte verwendet. Stattdessen werden die Objekte durch Sensorknoten ersetzt, die, nachdem sie der virtuelle Radarstrahl erfasst hat, aktiv Signale aussenden. Der Radarstrahl ist ebenfalls ein Funksignal, das die Sensorknoten triggert. Dieses Funksignal enthält Informationen über das physikalische Phänomen, das beobachtet werden soll, so dass die verteilten Sensorknoten nur noch mit „ja“ oder „nein“ antworten. Ein „ja“ bedeutet in diesem Fall, dass sie als Antwort ein Signal aussenden, bei „nein“ wird nichts ausgesendet. In Abbildung 2.10 ist das Szenario skizziert. Für das System ist es notwendig, dass zwischen „Radargerät“ und den Sensorknoten eine direkte Sichtverbindung besteht, damit eine genaue Schätzung der Laufzeiten möglich ist. Mit Hilfe dieser Laufzeiten werden dann die Positionen der Knoten bestimmt. Überlagerungen von Funksignalen entstehen, wenn Knoten gleichzeitig auf einen Trigger antworten, also gleichzeitig im Antennenbeam des „Radargerätes“ sind. Durch die Auswertung mehrerer Momentaufnahmen ist man in der Lage, die Positionen der antwortenden Funkknoten zu schätzen. Dies ist aus der Radartechnik als *synthetische Apertur* (SAR) bekannt. Die Überlagerung von Funksignalen wird hier allerdings nicht als Vorteil verstanden, sondern erhöht eher die Komplexität der Rekonstruktion. Die bekannten Methoden für SAR setzen allerdings voraus, dass die empfangenen Signale kohärent zu den ausgesandten Signalen sind, da sie ja als Reflektion von der gleichen Quelle stammen. Für die Sensorknoten, die die Reflektion nur imitieren, gilt dies nicht. Ihre lokalen Oszillatoren sind nicht aufeinander synchronisiert. Deshalb schlagen die Autoren eine Auswertung nur der Amplitude vor und leiten einen optimalen Maximum-Likelihood-Schätzer her. Hierbei liegt der Fokus jedoch etwas anders als in der vorliegenden Arbeit. Ziel ist es, die überlagerten Signale so auszuwerten, dass Positionen der Sensorknoten

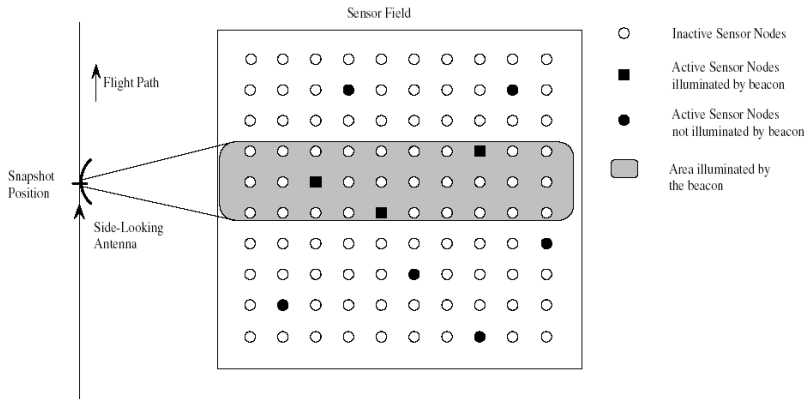


Abbildung 2.10. Ein mobiler Funkknoten sammelt Funksignale aus einem Sensornetz ein. Das System imitiert eine Radarstrahlreflektion; aus [8, Abbildung 1, Seite 2]

ermittelt werden können. Die Überlagerung wird also nicht als Vorteil verstanden, sondern eher als störender Nebeneffekt. Im Gegensatz dazu verfolgt die vorliegende Arbeit die Idee der Kooperation zwischen Knoten, um Überlagerungen bewusst vorteilhaft zu verwenden und kooperative Datenübertragung oder andere Signalisierung zu ermöglichen.

2.3 Bekannte Verfahren zur Zeitsynchronisation in Sensornetzwerken

Für alle Modelle von überlagerten Funksignalen ist Zeitsynchronisation eine *Notwendigkeit*, denn Überlagerung kann nur dann stattfinden, wenn Signale eine gewisse *Gleichzeitigkeit* aufweisen. Deshalb ist es wichtig, eine Zeitsynchronität im Netzwerk zu erreichen. Zeitsynchronisation ist eine Standardfragestellung in verteilten Systemen, deshalb gibt es auch bereits eine Vielzahl von Lösungen, wie Zeitsynchronisation in solchen Systemen durchgeführt werden kann. Für den Anwendungsfall dieser Arbeit gibt es jedoch besondere Anforderungen. Da die Zeitsynchronisation die Basis für die Überlagerung ist, so ist z.B. eine sehr genaue

Synchronisation nötig. Will man mit überlagerten Funksignalen z.B. Datenübertragung realisieren, so sind Synchronisationsgenauigkeiten nötig, die deutlich unterhalb der Symboldauer liegen. Typischerweise wird man hier Genauigkeiten im μs -Bereich anstreben.⁷

Die wohl bekanntesten Synchronisationsverfahren sind GPS [25] und das Network Time Protocol [57], das im Internet vorherrscht. In [19] und [9] werden Mechanismen zur Zeitsynchronisation in drei Gruppen klassifiziert:

Detektion von Resynchronisationseignissen. Hier wird die Zeit identifiziert, zu der eine Resynchronisation der verteilten Uhren nötig ist. Resynchronisationen können periodisch oder ereignisbasiert passieren. So könnte z.B. ein ausgewählter Knoten periodisch die verteilten Uhren mithilfe einer Nachricht resynchronisieren.

Schätzung von entfernten Uhren. Hier wird die lokale Zeit von entfernten Knoten geschätzt. Dies wird typischerweise durch Austausch von Paketen realisiert, die lokale Zeitstempel tragen. Je nach Kenntnis der Verzögerung durch den Transport der Pakete sind die Schätzungen besser oder schlechter.

Nachstellen von Uhren. Dies ist eine Kombination von den beiden voranstehenden Verfahren. Zum Moment einer Resynchronisation wird, basierend auf der Schätzung entfernter, lokaler Uhren, eine neue lokale Zeit gesetzt.

Typischerweise ist der Fokus einer Zeitsynchronisation, die maximale oder mittlere Abweichung aller lokaler Uhren von einer Referenzuhr zu minimieren. Hierfür gibt es in der Literatur eine große Anzahl verschiedener Ansätze. In den folgenden Unterkapiteln werden nun die bekanntesten Synchronisationsmechanismen für Sensornetzwerke vorgestellt, und zwar zunächst solche, die auf „normaler“ Datenübertragung basieren und ohne die Nutzung von überlagerten Funksignalen arbeiten. In Kapitel 2.3.4 werden dann noch zwei Verfahren vorgestellt, die Zeitsynchronisation mit Hilfe von überlagerten Funksignalen realisieren. Ein weiteres Verfahren wird in Abschnitt 5.3 aufgegriffen.

⁷Rechenbeispiel: Die Datenübertragung soll mit einer Baudrate von 100kbaud realisiert werden. Damit ist die Symboldauer $1/100000=10\mu s$. Um eine Gleichzeitigkeit von ausgesandten Signalen verschiedener Quellen zu erreichen, sind also Synchronisationsgenauigkeiten im Bereich $\leq 3\mu s$ nötig.

2.3.1 LTS

Das „Lightweight Time Synchronisation“-Protokoll [19], in 2003 von J. Greunen und J. Rabaey vorgestellt, verwendet paarweise Synchronisationen, die dann in Baumstrukturen auf Multihop-Netze ausgeweitet werden. Für die paarweise Synchronisation wird ein Paket ausgetauscht und dabei werden die Aussende- und Empfangszeitpunkte notiert. Nachdem ein Nachrichtenaustausch zwischen den Synchronisationspartnern zirkuliert hat, kann so die unbekannte Übertragungszeit herausgerechnet werden. Dies gelingt allerdings nur dann, wenn sie für alle Knoten identisch ist. Die Übertragungszeit setzt sich nach [19] aus

- Sendezeit
- Signallaufzeit
- Empfangszeit
- Kanalzugriffszeit

zusammen. Für die Übertragung mit Funkwellen und Szenarien von Sensornetzwerken ist die Signallaufzeit hierbei sicherlich nicht sehr relevant⁸. Darüber hinaus wird angenommen, dass alle anderen Zeiten gaußverteilt sind, speziell auch die Kanalzugriffszeit. Diese Annahme ist allerdings in der Standardliteratur keineswegs nachvollziehbar. Die Sende- und Empfangzeiten sind auch stark davon abhängig, auf welcher Schichtenebene die Empfangszeit nun gemessen wird. Sensorknoten, die mehrere Aufgaben gleichzeitig bearbeiten können, können hier keine Garantien im μs -Bereich geben. Die Autoren kommen zu der Erkenntnis, dass mit typischen Sensorknoten eine Genauigkeit von 100 μs mit einer Sicherheit von 99% erreicht werden kann. In den Simulationen erreichen die Autoren typische Synchronisationsgenauigkeiten im Bereich von mehreren 100 ms. Dies ist sicher nicht genug, um es als Basis für die Synchronisation von Knoten zu verwenden, die überlagerte Funksignale einsetzen. Die Synchronisationszeit wäre ein Vielfaches der Signallänge und damit völlig ungeeignet.

2.3.2 RBS

Die Synchronisation mit Hilfe von Referenzrundrufen (Reference broadcast synchronization) ist im Jahr 2002 von J. Elson, L. Girod und D. Estrin veröffentlicht worden [18]. Sie schlagen ein Synchronisationsverfahren vor, das eine hohe Genauigkeit (typischerweise $<5\mu s$) erzeugt und auf Rundruf auf der physikalischen

⁸Rechenbeispiel: Bei einer Entfernung von 100m ist die Lichtlaufzeit nur ca. 0.3 μs

Schicht beruht. Die Empfänger von Rundrufpaketen nutzen die Ankunftszeiten als Referenzen für einen Uhrenvergleich. Die einfachste Form könnte folgendermaßen aussehen:

- Ein Sender sendet ein Referenzpaket an zwei Empfänger.
- Jeder Empfänger notiert die Ankunftszeit in seiner lokalen Uhrzeit.
- Die Empfänger vergleichen anschließend ihre Empfangszeitpunkte.

Mit dieser Information können die Empfänger ihre lokalen Uhren in Abgleich bringen. Es ist keine weitere Information nötig. Es wird so jedoch nur eine *relative* Zeitsynchronisation erreicht, keine absolute, da von dem Sender keine Zeitangabe mitgeschickt wird. Als nächste Erweiterung schlagen die Autoren einen Multihop-Betrieb vor, der auch wieder nur auf dem Empfang von Synchronisationspaketen beruht. Abbildung 2.11 zeigt den Aufbau des Multihop-Verfahrens. Die Netzwerkknotten A, B, C, D senden jeweils Referenzpakete und die Knoten 1, 2,... 11 empfangen die Pakete. Die Knoten 4, 7, 8 und 9 empfangen Referenzpakete von unterschiedlichen Sendern und können so die relative Zeitsynchronisation zwischen den Bereichen vermitteln. Die Synchronisationen sind in allen Fällen auf Paketaustausch angewiesen. Optimale Koordination für die Ausbreitung einer Synchronisation wird dabei nicht angesprochen. Gerade in dichten Szenarien muss eine große Anzahl von Paketen ausgetauscht werden, was einen großen Aufwand bedeutet.

2.3.3 Mini-sync und Tiny-sync

M. Sichitui und C. Veerarittiphan haben 2003 Synchronisationsmethoden für drahtlose Sensornetzwerke veröffentlicht, die sie *tiny-sync* und *mini-sync* nannten [73]. Die Synchronisationsmethoden befassen sich hauptsächlich damit, den Drift zwischen lokalen Uhren zu erfassen. Dieser Drift entsteht aus der Variation zwischen den genauen Oszillationsfrequenzen der lokalen Quarze oder Schwingkreise. Die Autoren modellieren den Zusammenhang zwischen einer lokalen Uhrzeit t_i und der tatsächlichen Uhrzeit t als linear:

$$t_i(t) = a_i t + b_i \quad [73, \text{Formel}(1)] \quad (2.3)$$

Somit lässt sich der Zusammenhang zwischen zwei lokalen Uhren folgendermaßen beschreiben:

$$t_1(t) = a_{12} t_2(t) + b_{12} \quad [73, \text{Formel}(2)] \quad (2.4)$$

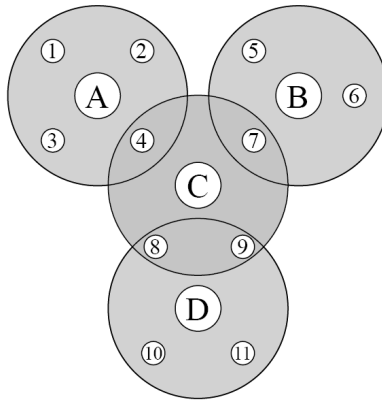


Abbildung 2.11. Relative Zeitsynchronisation mit *RBS* in einer Multihop-Topologie; aus [18, Abbildung 7a, Seite 14]

Wenn z.B. Knoten 1 und Knoten 2 ihre lokalen Zeiten abgleichen wollen, so wird ein Paket von 1 nach 2 geschickt und von Knoten 2 direkt wieder zurückgeschickt. Der Vorgang ist in Abbildung 2.12 skizziert. Da die kausale Reihenfolge der Zeitpunkte $t_0 < t_b < t_r$ ist, können Ungleichungen aufgestellt werden, mit denen obere und untere Schranken für die Unbekannten a_{ij} und b_{ij} gefunden werden. Je mehr solcher Pakete ausgetauscht werden, desto besser können die Bedingungen für die Schranken festgelegt werden. Hier wird nun zwischen *mini-sync* und *tiny-sync* unterschieden. Bei *tiny-sync* wird stets nur eine minimale Anzahl von Zeitwerten gehalten, um die Unbekannten durch Schranken zu beschreiben. Somit kann es vorkommen, dass während der Laufzeit Zeitwerte verworfen werden, mit deren Hilfe aber in der Zukunft eine bessere Schätzung möglich geworden wäre. In der Erweiterung *Mini-sync* werden mehr als nur die nötigsten Zeitpunkte gespeichert. In einer empirischen Untersuchung zeigen die Autoren, dass typischerweise 40 solcher Werte ausreichen, um eine nahezu optimale Lösung zu erhalten. Die erreichte Synchronität lag in den Simulationen bei durchschnittlich 3ms für Singlehop-Netze und 30ms für Multihop-Netze. Diese Genauigkeit ist für die Anwendung zur Synchronisation für überlagerte Funksignale nicht ausreichend. Die Synchronisation muss deutlich unterhalb der Symboldauer liegen.

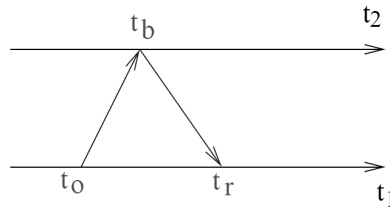


Abbildung 2.12. Ein Paket wird von Knoten 1 nach 2 geschickt und direkt wieder zurückgeschickt. Die Zeitpunkte t_0 , t_b und t_r werden notiert.; aus [73, Abbildung 1, Seite 3]

Aus diesen Synchronisationsverfahren würden Symbolauern von einigen 10ms resultieren, was Datenraten von nur einigen 10Byte/sec entspricht.

2.3.4 Überlagerte Funksignale zur Zeitsynchronisation

In jüngster Vergangenheit wurden auch überlagerte Funksignale zur Zeitsynchronisation verwendet. Die erste hier vorzustellende Arbeit ist von Y. W Hong und A. Scaglione [28]. Sie schlagen ein System gekoppelter Oszillatoren vor, das von den biologischen Vorgängen gekoppelter Neuronen oder Schwarmverhalten motiviert wird. Hierfür werden Ultrabreitbandimpulse eingesetzt. Die Teilnehmer dieses Synchronisationsverfahrens senden also periodische Breitbandimpulse aus. Das Verfahren ist speziell für Multihop-Topologien geeignet, sehr einfach aufgebaut und deshalb sehr gut für Sensornetzwerke geeignet. Ziel des Systems ist es *nicht*, eine Synchronisation global vorzugeben, und diese in dem Netzwerk zu etablieren, sondern vielmehr mit Hilfe von lokaler Adaption eine gemeinsame Schwingung zu finden, die die größte Übereinstimmung erzeugt. Als Systemvoraussetzung wird festgestellt, dass die Zeitdauer der Breitbandimpulse deutlich kürzer sein muss als die Periode der Oszillation. Abbildung 2.13 zeigt das Konvergenzverhalten für eine Simulation von 40 Netzwerkknoten. Die Autoren schlagen zusätzlich eine Detektionsvorschrift und eine Synchronisationsstrategie vor. Die genauen Parameter werden in der genannten Veröffentlichung allerdings nicht festgelegt. Die Autoren verwenden und präzisieren das vorgestellte Verfahren in weiteren Veröffentlichungen [29], [30]. Hier geht es stets um die Eigenschaft, ei-

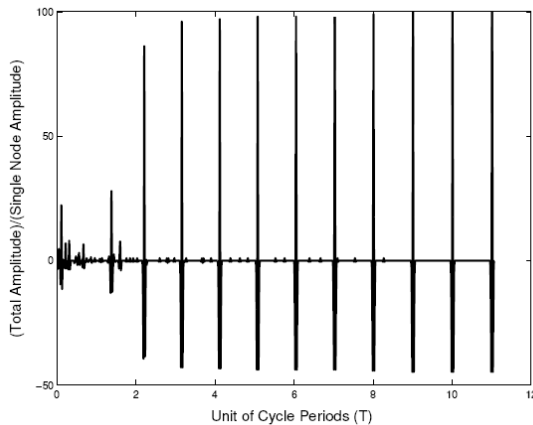


Abbildung 2.13. Die Amplitude des Empfangssignals in weiter Entfernung normalisiert auf die Leistung nur eines Senders; aus [28, Abbildung 3]

ne lokale Information durch die gekoppelten Oszillatoren zu verteilen, also eine Informationsverteilung durch Rundruf.

Die detailliertesten Arbeiten, die überlagerte Funksignale zur Synchronisation analysieren, sind von An-swol Hu und S. Servetto veröffentlicht worden [34], [33]. Die Autoren befassen sich in den umfangreichen Veröffentlichungen mit drei Themen:

- ein probabilistisches Modell zur Beschreibung
- *spatial averaging* als Konzept zur Zeitsynchronisation
- Modelle für verschiedene Ausbreitungen, Ausbreitungszeiten

In Abbildung 2.14 sieht man die zentrale Idee, die in den Arbeiten der Autoren verfolgt wird. Sie bezeichnen sie mit *spatial averaging*. Ein Signal — in diesem Fall ein Rechteckimpuls — ist das Basissignal für die Synchronisation. Jeder Teilnehmer sendet ein solches Signal (oder einen Impuls) aus. Falls nun die Zeitsynchronisation zwischen den Knoten mittelwertfrei ist, so erreicht man durch den zentralen Grenzwertsatz eine Gaußverteilung für den überlagerten Fall. Die Flanke aus Abbildung 2.14 (linke Seite) wird durch vielfache Überlagerung zu einer Gaußflanke auf der rechten Seite. Die Mittelung („*spatial averaging*“) kann

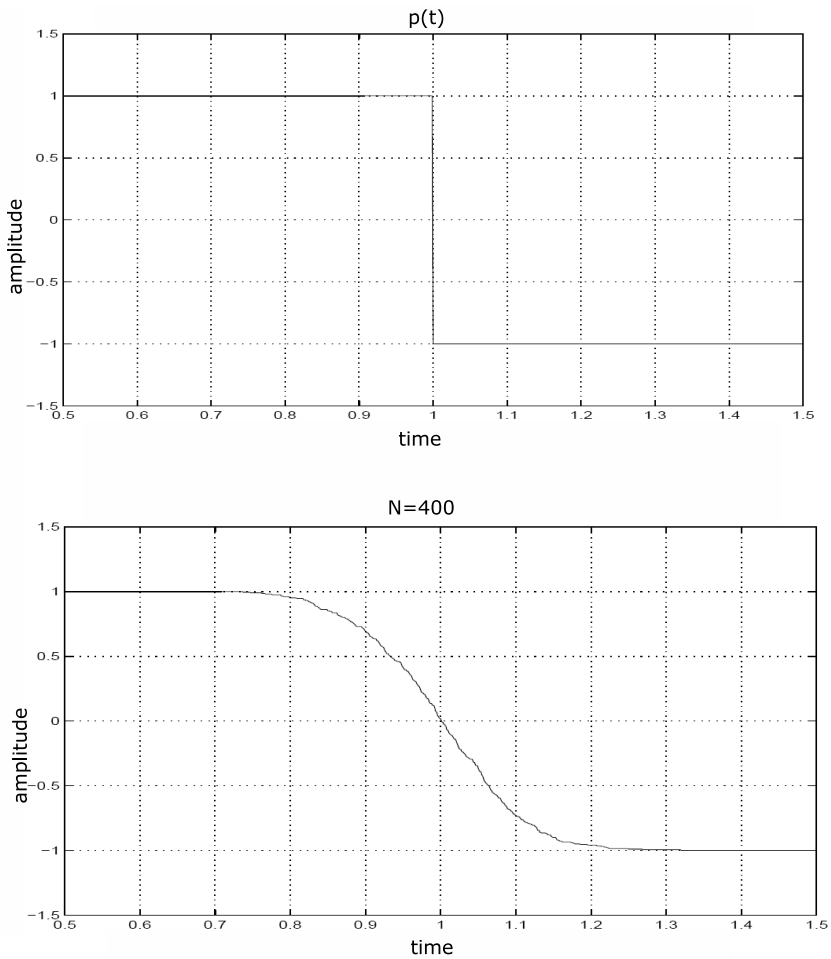


Abbildung 2.14. Die Überlagerung von 400 Rechtecksignalen (unten). Sie stammen von 400 Knoten, die eine leichte Zeitverschiebung haben, die symmetrisch und offsetfrei um die Flanke des Rechtecks verteilt sind; aus [34, Abbildung 1]

nun verwendet werden, um den Nullpunkt oder den Mittelwert aller einzelnen Signale zu bestimmen. In Abbildung 2.14 ist zu sehen, dass der Mittelwert (halbe Signalhöhe) unverändert zur Zeit 1 auftritt. Die Überlagerung der Signale war also erfolgreich. Die Theorie wird dann auf eine Klasse von Impulsen $p(t)$ erweitert, die punktsymmetrisch sind, also $p(t) = q(t)$ und $p(-t) = -q(t)$, somit auch $p(0) = 0$. Dies vereinfacht das Finden der Synchronisationszeitpunkte. Sie sind dann die Nulldurchgänge der summierten Signale. Abbildung 2.15 zeigt das Ergebnis der Überlagerung solcher Impulse.

Der entscheidende Unterschied der vorliegenden Arbeit besteht darin, dass die Zielplattform „normale“ Bandpasssignale verwenden soll. Der Vorschlag der Autoren, der in Abbildungen 2.14 und 2.15 dargestellt ist, geht davon aus, dass das ausgesandte Signal ein punktsymmetrisches Rechtecksignal oder Puls ist. Die Betrachtungen gelten daher direkt für die Signale im Kanal. Somit müsste ein System in der Lage sein, diese Signale aus den Abbildungen 2.14 und 2.15 direkt im Bandpass- oder Breitbandbereich zu erzeugen. Dies erfordert wiederum einen erheblichen Aufwand in der Signalerzeugung. Die Rechtecksignale sind z.B. überhaupt nicht als Bandpasssignal gedacht und sind so gar nicht im Bandpassbereich darstellbar. Eine einfache Verschiebung in den Bandpassbereich durch Aufmodulation auf einen Träger ist nicht möglich, da sich dann die Überlagerung völlig verändert. Das vorgeschlagene Verfahren ist zwar prinzipiell technisch realisierbar und — obwohl es nie in Hardware implementiert wurde — höchstwahrscheinlich auch in der Praxis funktional. Problematisch sind dabei jedoch zwei Punkte: Der Bandbreitenbedarf ist nicht diskutiert worden. Das Verfahren der Rechteckimpulse funktioniert nur ohne Verschiebung in den Bandpassbereich und kann so keine Lizenzierung finden; die Erzeugung der Breitbandsignale ist technisch einfacher, erfordert aber auch eine hochgenaue Startzeitregelung, um die Überlagerung der Signale zu erreichen. Darüberhinaus soll die Zielplattform der vorliegenden Arbeit nicht mit Breitbandsignalen arbeiten, und so ist auch dieser Vorschlag nicht als Lösung geeignet.

2.4 Thematische Einordnung und Positionierung dieser Arbeit

Wie im vorangegangenen Kapitel klargestellt, sind überlagerte Funksignale Gegenstand aktueller, internationaler Forschung. Deshalb ist es wichtig, die vorliegende Arbeit in die internationalen Aktivitäten einzuordnen und den Fokus klar

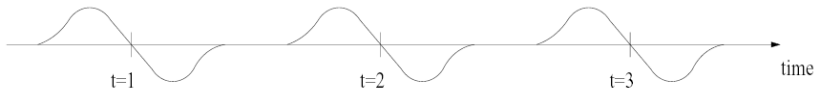


Abbildung 2.15. Überlagerung punktsymmetrischer Impulse. Der Nulldurchgang ist der Synchronisationszeitpunkt; aus [34, Abbildung 5]

herauszustellen. Der Fokus dieser Arbeit wurde im Analysekapitel (2.1) schon teilweise festgelegt und soll nur weiter präzisiert werden.

Aufgrund der harten Anforderungen für die Zielplattform sind viele bekannte Mechanismen nicht möglich, die in den vorangegangenen Kapiteln vorgestellt wurden. Im Speziellen geht es um zwei Hauptvoraussetzungen, nämlich

- „line of sight“ (LOS) und
- komplexwertige Signalverarbeitung (digital signal processing DSP).

LOS bedeutet, dass das System voraussetzt, dass es eine direkte Sichtverbindung zwischen Quelle und Ziel gibt. Diese Sichtverbindung impliziert auch, dass zwischen Sendern und Empfänger eine bidirektionale Verbindung besteht, also alle Quellen und das Ziel in direkter, individueller Funkverbindung stehen. In den meisten Fällen werden auf diesen Funkverbindungen feste Träger verschickt, damit die Phasen und Frequenzvorsätze der Funksignale der Kommunikationspartner erfasst werden können. Diese direkte Verbindung zwischen Quellen und Ziel ist also eine Voraussetzung, um Phasen- und Frequenzsynchronisationen zu bewerkstelligen. Das Motivationsbeispiel in Kapitel 1.3 hat schon gezeigt, dass für die Szenarien und Anwendungsfälle dieser Arbeit der direkte Sichtkontakt zwischen Quelle und Ziel nicht vorausgesetzt werden kann. Somit ist *keines* der bekannten Verfahren anwendbar, das eine Phasen- und Frequenzsynchronisation etablieren soll. Deshalb soll Phasen- und Frequenzsynchronisation auch nicht als Voraussetzung oder Anforderung an die Beiträge dieser Arbeit gestellt werden. Überlagerte Funksignale sollen *ohne* diese zwei Synchronisationen auskommen. Dies impliziert automatisch, dass keine Kohärenz (also eine Phasen-, Zeit- und Frequenzsynchronität) erzeugt werden kann.

Der nächste Punkt ist, dass für die Zielplattform nicht vorausgesetzt werden kann, dass komplexwertige Signalverarbeitung möglich ist. Wie in Kapitel 2.1

beschrieben, haben die Sensorknoten sehr einfache Frontends und verarbeiten die Basisbandsignale mit einfachen analogen Filtern und Regelungen. Oft ist selbst der Zugriff auf das Basisband und ein Eingriff in die Modulation nicht möglich.

Wie in Kapitel 2.3 schon analysiert, muss die Zeitsynchronisation im Bereich von wenigen μs liegen, wenn man überlagerte Funksignale ermöglichen will. In den vorgestellten Verfahren, die paketorientiert sind, wird diese Genauigkeit nur von RBS erreicht. Verfahren zur Zeitsynchronisation, die selbst überlagerte Funksignale verwenden, erreichen vermutlich höhere Synchronisationsgenauigkeit. Die Arbeiten sind jedoch noch rein theoretischer Natur, so dass nicht einmal Simulationen vorliegen, die einen quantitativen Vergleich ermöglichen würden. Darüber hinaus stellen sie hohe Anforderungen an die Leistungsfähigkeit der eingesetzten Hardware. Der Einsatz von Ultrabreitbandsignalen ist nicht nur technisch, sondern auch aus Sicht der Lizenzierung noch nicht praktikabel. Es ist also auch im Bereich Zeitsynchronisation weitere Forschung nötig, um der Implementierung von überlagerten Funksignalen näher kommen zu können.

Verfahren	Voraussetzungen		Synchron.			Kanalmode	
	LOS	DSP	f_c	φ	t	AWGN	Mehrwege
Hillenbrand et al.		x			x	x	x
Yung-Szu Tu et al.	x	x	x	x	x	x	
Barriac et al.	x	x	x	x	x	x	
B. Anatha.	x	(x)			x	x	
D. Brown	x	x	x	x		x	(x)
Scaglione et al.		x		x	x	x	x
Servetto et al.	x	x	x	x	x	x	
diese Arbeit					x	x	x

Tabelle 2.1. Charakterisierung anderer Arbeiten auf dem Gebiet der überlagerten Funksignale

Tabelle 2.1 zeigt die in den vorangegangenen Kapiteln vorgestellten Arbeiten in einer Übersicht. Die Arbeiten werden nach verschiedenen Kriterien charakterisiert. Unterscheidungen finden sich zunächst in den Systemvoraussetzungen. Hier werden den Verfahren zwei besonders wichtige Eigenschaften zugeordnet, nämlich der direkte Funkkontakt („LOS“) und die Fähigkeit, komplexwertige (De-)Modulationen zu verarbeiten („DSP“). Als Nächstes werden die nötigen Synchronisationen unterschieden. Die vorliegende Arbeit ist die einzige, die mit minima-

len Systemvoraussetzungen und nur einer Zeitsynchronisation mit überlagerten Funksignalen arbeiten kann. Außerdem ist sie eine der wenigen Arbeiten, die für das Kanalmodell auch Mehrwegeausbreitung der Signale betrachtet.

Abbildung 2.1 soll nun erweitert werden, um die relevanten Arbeiten im Forschungsbereich und auch die vorliegende Arbeit einzuordnen. Dies ist in Abbildung 2.16 zu sehen. Die vorliegende Arbeit befasst sich also mit Funksignalen, die

- nicht orthogonal,
- kontrolliert,
- im Bandpassbereich und
- nicht kohärent

überlagert werden und dabei auch noch

- skalar (oder rein analog)

verarbeitet werden. Auf Basis dieser Feststellungen lässt sich nun definieren, was unter überlagerten Funksignalen in der vorliegenden Arbeit verstanden wird:

Definition 2.1 (Überlagerte Funksignale) *Überlagerte Funksignale sind Signale im Bandpassbereich, die zeitgleich ausgesendet werden und sich so im Funkkanal nicht orthogonal inkohärent überlagern.*

2.5 Zentrale Beiträge

Um überlagerte Funksignale auf Sensorknoten einzusetzen sind noch Verbesserungen und Lösungen zu erforschen. Die strikten Einschränkungen und Anforderungen, die mit low-cost Hardware einhergehen, führen dazu, dass herkömmliche Ansätze nicht verwendbar sind. Die vorliegende Arbeit will hier Beiträge leisten, um überlagerte Funksignale für Sensorknoten realisierbar zu machen. Es steht in dieser Arbeit die Realisierung und Anwendung im Vordergrund. Theoretische Beiträge werden hier nur als Mittel zum Zweck verstanden. Die Anwendungen überlagerter Funksignale stehen klar im Vordergrund. Im Einzelnen umfasst diese Arbeit folgende wissenschaftliche Beiträge:

Ein Modulationsverfahren, das mit den Einschränkungen für die Zielplattform auskommt. Das Übertragungsverfahren stellt sich auch Signalformung und

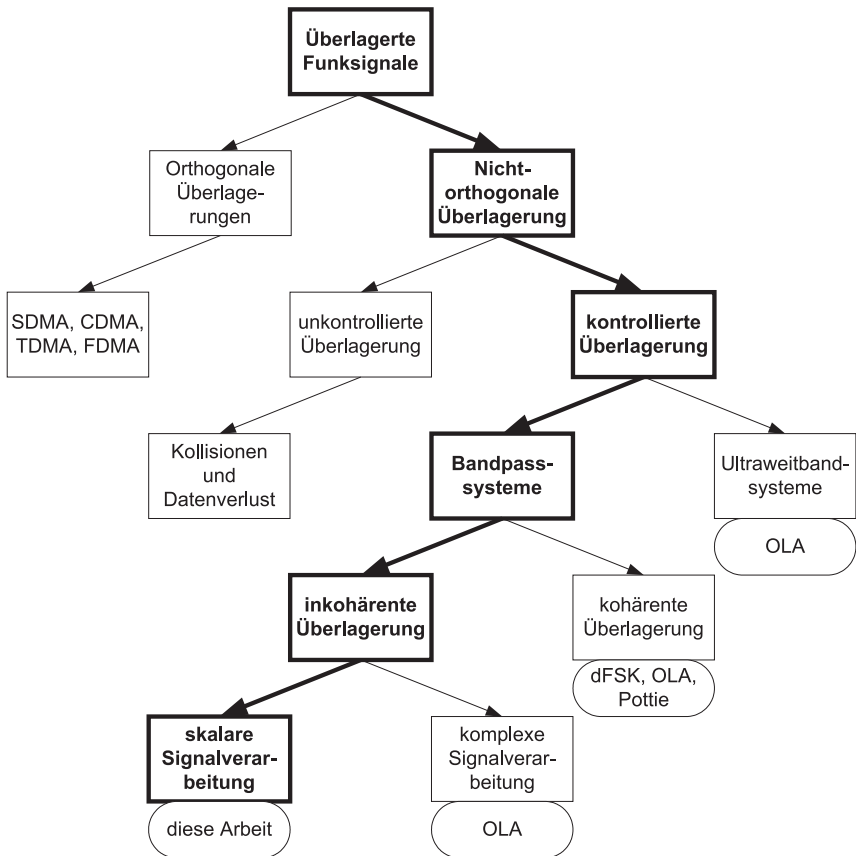


Abbildung 2.16. Thematische Einordnung dieser Arbeit: Es geht um Funksignale, die nicht orthogonal, kontrolliert und nicht kohärent im Bandpassbereich überlagert werden und nur skalar verarbeitet werden

Detektion zusammen. Hier wird ein neues, mehrstufiges Modulationsverfahren vorgeschlagen, dass für Überlagerungen geeignet ist und sich als *ESK* in die Liste von OOK, ASK, FSK, PSK etc. einreicht.

Ein Zeitsynchronisationsverfahren, das sehr genaue Synchronisationen erreicht. Hohe Zeitsynchronisation ist als Basis für überlagerte Funksignale unerlässlich. Das vorgestellte Verfahren basiert selbst auch auf Überlagerung und kann sowohl in Stern- als auch in Multihop-Topologien eingesetzt werden. Es erreicht Synchronisationen $\leq 4\mu s$ zwischen allen Teilnehmern eines Netzes.

Ein Datenübertragungsverfahren, welches überlagerte Funksignale nutzt. Es ist speziell für die **N zu 1** Datenkommunikation geeignet, da es eine Prozesszeit garantiert, die *unabhängig* von der Anzahl N der Teilnehmer ist. Es ist das einzige bekannte Protokoll, das in diesem Sinne *skalierungsfrei* ist. Es ermöglicht extrem schnelle Datenübertragung und kann Datenfusion bereits auf der physikalischen Schicht durchführen.

2.6 Systemaufbau und Architektur

In Abbildung 2.17 wird der Systemaufbau dargestellt. Das zu betrachtende System besteht aus Sensorknoten und dem Übertragungsmedium. Die Sensorknoten stehen typischerweise über das Funkmedium in direktem Kontakt zu ihren Nachbarn. Sie sind besonders einfach aufgebaut und können deshalb keine komplexe Signalverarbeitung leisten. Für die Überlagerung auf dem Medium wird — wie im vorangegangenen Kapitel analysiert — auf Träger- und Phasensynchronisation verzichtet. Dennoch sollen überlagerte Funksignale Beiträge zu Zeitsynchronisation, Datenübertragung und -fusion, sowie Zuverlässigkeit und effizienter Kanalnutzung leisten.

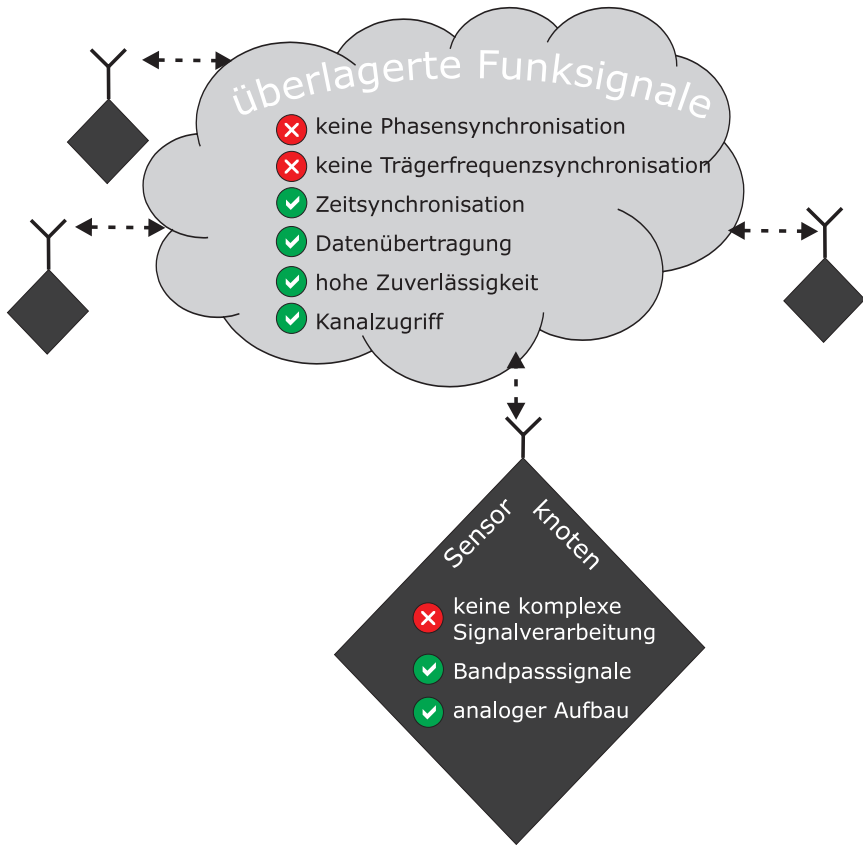


Abbildung 2.17. Architektur der Kommunikation und der Transceiver der Sensorknoten für überlagerte Funksignale

3 Übertragungsmodelle

Um überlagerte Funksignale technisch nutzen zu können, ist es notwendig, spezielle Signalformungen in den Sendern durchzuführen. Durch die Gleichzeitigkeit auf dem Kanal entstehen Überlagerungsmuster, die für die Erkennung im Empfänger schädlich und vorteilhaft sein können. Die Einflüsse der Überlagerungen auf die Signale müssen also bewusst eingesetzt werden, um den größtmöglichen Vorteil daraus zu ziehen. Dieses Kapitel befasst sich mit dieser Signalformung in den Sendern und gleichzeitig mit der passenden Erkennung (Detektion) im Empfänger. Dabei sind insbesondere die harten Anforderungen an die Signalverarbeitung durch die leistungsschwache Hardware eine zu überwindende Hürde. Im Speziellen werden nun zwei verschiedene Übertragungsszenarien betrachtet, für die eine vollständige Modellbildung von Signalformung über Einflüsse durch die Übertragung und Überlagerung bis hin zur Erkennung durchgeführt wird. Es wird ein neues, passendes Signalformungsverfahren (Modulationsverfahren) analytisch hergeleitet, das für die leistungsschwache Hardware geeignet ist und eine optimale Übertragung mit überlagerten Funksignalen ermöglicht.

Die Systemarchitektur aus Kapitel 2.6 wird nun unter verschiedenen, speziellen Voraussetzungen präzisiert und Signalformungen und Detektionsvorschriften hergeleitet. Diese speziellen Voraussetzungen werden sich vor allem aus den Annahmen zur Synchronisation ergeben. Auch zum Kanal werden verschiedene Annahmen getroffen werden. Das Ziel ist es, ein System zu erhalten, das den harten Anforderungen aus Kapitel 2.1 gerecht wird.

Es wird nun das Systemmodell für eine Überlagerung von Signalen hergeleitet, das als Basis weiterer Betrachtungen herangezogen wird. Die Architektur des Systems besteht aus N sendenden Knoten, die mit l ; $1 \leq l \leq N$ indiziert sind. Diese sollen zeitgleich dasselbe komplexe Symbol $\underline{s} = a + jb$ übertragen. Sie tun dies mit einer komplexen Modulation auf ihrem lokalen Träger ω_l . Zwar sollen nach den Voraussetzungen aus Kapitel 2.1 solche komplexen Modulationen vermieden werden; sie sind aber in der allgemeinen Modellbildung hilfreich. Es wird an späterer Stelle dann eine Spezialisierung und Vereinfachung auf eine nicht-komplexe Modulation erfolgen. Die lokalen Trägerfrequenzen ω_l weichen von der Trägerfrequenz ω_c des Empfängers ab, so dass gilt: $\omega_l = \omega_c + \Delta\omega_l$. Auch die Phasen der Sender und des Empfängers sind prinzipiell nicht aufeinander abgestimmt, so dass eine lokale Phasenverschiebung φ_l existiert. Genauso bleibt auch eine Differenz in der Zeitsynchronisation der Symbole gegenüber dem Empfänger von τ_l übrig. Das Symbol wird mit einer Fensterfunktion $w(t)$ im Zeitbereich gewichtet. Zusammenfassend werden nun die folgenden Bezeichner eingeführt:

$\underline{s} = a - jb$ das Symbol, das von der Sendern verschickt wird

N die Anzahl der Sender

l der Laufindex über die Anzahl N von Sendern

ω_l die lokale Trägerfrequenz eines Senders

ω_c die lokale Trägerfrequenz des Empfängers

φ_l die lokale Phasenverschiebung des Trägers eines Senders gegen den Empfänger

τ_l die Symbolstartzeitverschiebung eines Senders gegen die Startzeit des Empfängers

$t_l(t)$ das ausgesandte Signal des l -ten Senders

$\underline{t}_l(t)$ das komplexe Signal als Ergebnis der lokalen, komplexen Modulation

$w(t)$ die Fensterfunktion, die die Signalformung des ausgesandten Symbols festlegt

\underline{n} das Rauschen im Empfänger (als gaußisch angenommen)

$\underline{r}_l(t)$ der Teil des Empfangssignals, das vom Sender l zum gesamten Empfangssignal beigetragen wird

$\underline{r}(t)$ das Empfangssignal als Summe aller Signale ($= \sum \underline{r}_l(t)$)
 $\underline{r}_{l,BP}(t)$ das Empfangssignal eines Senders in der Bandpasslage (vor Demodulation)
 $\underline{g}_l(t)$ die Impulsantwort des Kanals von Sender l zum Empfänger
 I, Q die Inphase- und Quadraturkomponenten der Signale (Kosinus- und Sinusanteile)
 j die imaginäre Einheit

Mit Hilfe dieser Definitionen ist es nun möglich, das Signal zu formulieren, das einer der N Sender abstrahlt. Für die Herleitung wird komplexe Signalübertragungstheorie angewendet [68]. Bei der Modulation wird das Symbol mit der Fensterfunktion gewichtet und mit der komplexen, lokalen Trägerschwingung moduliert:

$$\begin{aligned}
 \underline{t}_l(t) &= \underline{S} \cdot w(t - \tau_l) \cdot e^{j(\omega_l t + \varphi_l)} \\
 &= w(t - \tau_l) [m(t) + j\hat{m}(t)] \\
 \text{mit } m(t) &= a \cos(\omega_l t + \varphi_l) - b \sin(\omega_l t + \varphi_l) \\
 \text{und } \hat{m}(t) &= a \sin(\omega_l t + \varphi_l) + b \cos(\omega_l t + \varphi_l)
 \end{aligned} \tag{3.1}$$

Übertragen wird hiervon nur der Realteil:

$$\begin{aligned}
 t_l(t) &= \Re\{\underline{t}_l(t)\} \\
 &= w(t - \tau_l) m(t) \\
 &= w(t - \tau_l) [a \cos(\omega_l t + \varphi_l) - b \sin(\omega_l t + \varphi_l)]
 \end{aligned} \tag{3.2}$$

Im Kanal werden nun die von den N Sendern gleichzeitig ausgeschickten Signale $t_l(t)$ durch Überlagerung addiert. Der Empfänger demoduliert dieses Gemisch, wie in Abbildung 3.1 zu sehen. Das Signal nach der Demodulation und Filterung in die Tiefpasslage wird als \underline{r} bezeichnet. Dieses setzt sich aus den Teilbeiträgen $\underline{r}_l(t)$ und dem Rauschen zusammen:

$$\underline{r} = \sum \underline{r}_l + \underline{n}$$

Zur Berechnung eines Teilbeitrages \underline{r}_l müssen der Kanal und die Demodulation berücksichtigt werden. Bei der Demodulation muss der fehlende Imaginärteil von Ausdruck $\underline{t}_l(t)$ wieder hergestellt werden. Er kann durch Hilbertransformation

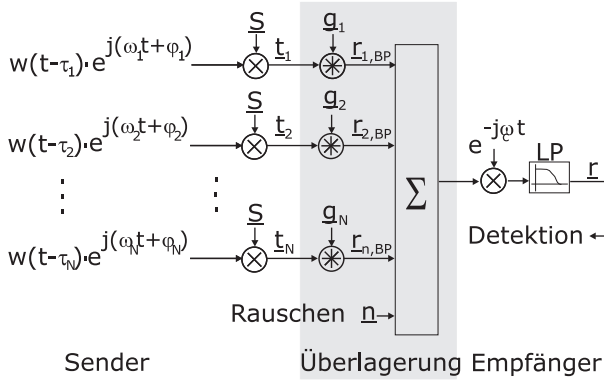


Abbildung 3.1. Systemmodell für überlagerte Funksignale

aus dem Realteil $t_l(t)$ gewonnen werden. ($\hat{m}(t) = \mathcal{H}\{m(t)\}$). Somit liegt dem Empfänger wieder das vollständige komplexe Signal $\underline{t}_l(t)$ vor:

$$\underline{r}_l(t) = \underline{t}_l(t) \star \underline{g}_l(t) \cdot e^{-j\omega_c t} \quad (3.3)$$

$$(3.4)$$

mit (3.1) ergibt sich:

$$\begin{aligned} \underline{r}_l(t) &= \underline{S} \cdot w(t - \tau_l) \cdot e^{j(\omega_l t + \phi_l)} \star \underline{g}_l(t) \cdot e^{-j\omega_c t} \\ \underline{r}_l(t) &= \underline{S} \cdot w(t - \tau_l) \cdot e^{j(\Delta\omega_l t + \phi_l)} \star \underline{g}_l(t) \end{aligned} \quad (3.5)$$

Ausgeschrieben ergibt sich:

$$\begin{aligned} \underline{r}_l(t) &= w(t - \tau_l) \underbrace{\{a \cos(\Delta\omega_l t + \phi_l) - b \sin(\Delta\omega_l t + \phi_l)\}}_{I\text{-component}} \\ &\quad + j \cdot \underbrace{\{b \cos(\Delta\omega_l t + \phi_l) + a \sin(\Delta\omega_l t + \phi_l)\}}_{Q\text{-component}} \star \underline{g}_l(t) \\ \underline{r}_l(t) &= w(t - \tau_l) \sqrt{a^2 + b^2} \cdot e^{+j(\Delta\omega_l t + \phi_l + \angle \underline{S})} \star \underline{g}_l(t) \end{aligned} \quad (3.6)$$

Ausgehend von dieser Formel (3.6) werden im Laufe dieser Arbeit verschiedene Systemvoraussetzungen für das Modell getroffen und daraus spezielle Darstellun-

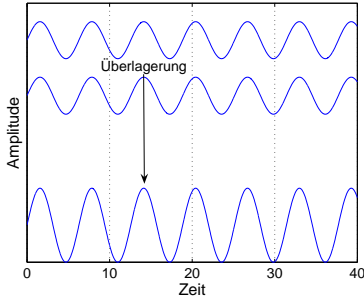


Abbildung 3.2. Kohärente Überlagerung führt zu konstruktiver Interferenz

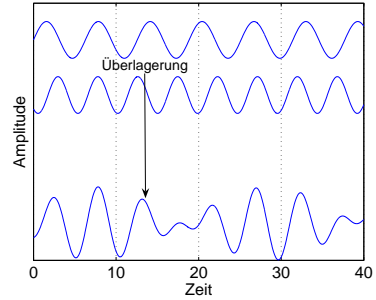


Abbildung 3.3. Inkohärente Überlagerung führt zu Schwebungen

gen erarbeitet. Im Idealfall wären die Sender auf den Empfänger exakt synchronisiert ($\tau_l = 0, \varphi_l = 0, \Delta\omega_l = 0$), die Kanäle $\underline{g}_l = \delta(t)$ und kein Rauschen vorhanden ($\underline{n} = 0$). Dann kontrahiert Ausdruck (3.6) zu:

$$\underline{r}_l(t) = \sqrt{a^2 + b^2} e^{j\angle \underline{S}} w(t) = \underline{S} w(t)$$

und das versendete Symbol \underline{S} wäre verlustfrei rückgewonnen. Dieser Fall ist allerdings in dieser Arbeit nicht relevant, da solche idealen Voraussetzungen nicht erfüllt sind. Stattdessen kann man in Formel (3.6) die typische Eigenschaft von überlagerten Funksignalen erkennen. Da der Empfänger seine lokale Frequenz ω_c nicht auf die des Senders (ω_l) synchronisiert hat, bleibt eine *Restoszillation* $\Delta\omega_l$ im Empfänger übrig.

Es soll nun noch einmal der Begriff der *nicht kohärenten Überlagerung* genauer definiert werden. Unter einer nicht kohärenten Überlagerung wird in dieser Arbeit eine Überlagerung verstanden, die nicht in Phase φ_l und Trägerfrequenz ω_l synchronisiert ist. Abbildungen 3.2 und 3.3 zeigen den Unterschied zwischen kohärenter und inkohärenter Überlagerung. In den Graphiken sind jeweils im oberen Bereich zwei Signale aus unterschiedlichen Quellen dargestellt. Das untere Signal ist das Summensignal. Für den kohärenten Fall sind die beiden Quellsignale in Phase und Frequenz gleich, und somit kommt es zu einer positiven Überlagerung oder *konstruktiven Interferenz*. In inkohärenten Fall sind Frequenz und Phase der Ursprungssignale nicht aufeinander abgestimmt. Das Summensignal enthält eine

Schwebung, und die Überlagerung oder Interferenz ist nicht zu allen Zeiten konstruktiv. Für die vorliegende Arbeit soll davon ausgegangen werden, dass sich die Funksignale inkohärent überlagern. Das liegt vor allem daran, dass das Übertragungssystem (bestehend aus Sendern, Kanal und Empfängern) nicht in der Lage ist, alle Signale phasenrichtig zur Deckung zu bringen. Es wird daher für den weiteren Verlauf der Arbeit angenommen, dass die Phasenwinkel φ_l gleichverteilt auf $[0, 2\pi[$ sind, und die Trägerfrequenzen ω_l gleichverteilt auf $[\omega_{min}, \omega_{max}]$. Die $\omega_{min}, \omega_{max}$ berechnen sich im System aus der Spezifikation der Quarze. So würde ein 2.4GHz Quarz mit 20ppm zu $[2.4GHz - 48kHz, 2.4GHz + 48kHz]$ führen. In den folgenden Kapiteln werden nun die Bereiche, die die Übertragung betreffen, analysiert und Signalmodelle für verschiedene Fälle hergeleitet:

Kooperatives Senden. Dieser Fall entspricht einem Szenario, wie es im Motivationsbeispiel in Kapitel 1.3 vorgestellt wurde. Eine Gruppe von Knoten sendet zu einem entfernten Ziel. Für das Kanalmodell wird hier ein *Flat-Fading*-Kanal angenommen. Die Dämpfungen der Signale von den Sendern zum Empfänger werden als annähernd gleich angenommen, da die Entfernung zwischen den Knoten gering gegenüber der Entfernung zum Ziel ist. Somit ergeben sich die Kanäle zu $g_l(t) = \underline{g}(t) = \alpha e^{j\beta} \delta(t)$. In einem solchen Szenario wird das Ziel verfolgt, mit der Überlagerung der Signale eine höhere Gesamtleistung zu erzielen, um so den entfernten Empfänger zu erreichen. Diese Herleitung und Signalformung wurde in [40] veröffentlicht.

Verteiltes Senden. Dieser Fall entspricht einem Szenario, in dem eine Gruppe Knoten Signale überlagert, die von Knoten der gleichen Gruppe empfangen werden. Die Entfernungen zu den Empfängern sind nicht gleich und die Kanäle können durch Abschattung und Mobilität beeinflusst werden. Als Kanalmodell wird hier der *rayleigh*-Kanal angenommen.

Detektion. Für die Detektion kann aufgrund der Einfachheit der Transceiver nur die Amplitude ausgewertet werden. Darüber hinaus sind auch keine hohen Abstraten für diese Amplitudenauswertung möglich. Um sogar einen analogen Detektor zu realisieren, ist es auch eine Option, eine Integration dieser Amplitude durchzuführen und die Detektion auf die *Energie* auszurichten.

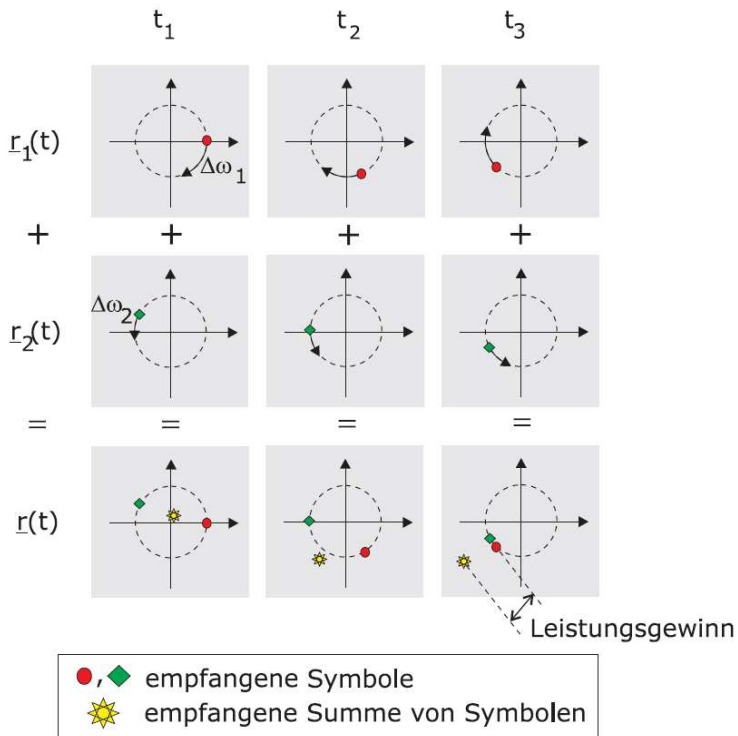


Abbildung 3.4. Rotation und Überlagerung im nicht-kohärenten Empfänger

3.1 Kooperatives Senden

Um ein besseres Bild von den Signalverläufen im Empfänger zu bekommen, dient Abbildung 3.4 als Grundlage. Die neun grau unterlegten Einzelbilder sind Momentaufnahmen von komplexen Signalmäumen. Die Markierungen stellen Momentaufnahmen von den Signalen dar. In diesem Beispiel wird ein Symbol von zwei Sendern ($N = 2$) gleichzeitig geschickt. Sie haben — bezogen auf den Empfänger — unterschiedliche Phasenverschiebungen φ_l ; $l = 1, 2$ und Restoszillationen $\Delta\omega_l$. Das Symbol wird konstant über die gesamte abgebildete Zeit (t_1 bis t_3) versendet. Die Restoszillationen führen dazu, dass die komplexen Symbole der Sender im Empfänger in der komplexen Signalebene rotieren. Dies ist mit Pfeilen angedeutet. Je nach Frequenzunterschied $\Delta\omega_l$ zwischen Sender und Empfänger ist diese Rotation unterschiedlich schnell und kann auch unterschiedliche Richtungen haben (je nachdem, ob Sender- oder Empfängerfrequenz größer ist). Der Empfänger sieht jedoch nicht die einzelnen, von den Sendern beigetragenen rotierenden Symbole, sondern lediglich die Summe dieser. Dies ist mit dem Stern in Abbildung 3.4 dargestellt. Je nach Lage der einzelnen Symbole kann dies in der Summe zu einer Auslöschung (t_1) oder zu einem Leistungsgewinn (t_3) führen.

3.1.1 Signalstatistiken für nicht kohärente Überlagerung

Um die Statistik der überlagerten Signale herzuleiten, kann man die Theorie des *Rayleigh-Fading*-Kanals anwenden. In dem Rayleighmodell wird ein Symbol (hier \underline{S}) übertragen, und den Empfänger erreichen mehrere Kopien dieses Signals mit gleicher Leistung, aber gleichverteilter Phase. Im Rayleighmodell wird insgesamt keine Leistungsänderung modelliert und angenommen, dass die Kopien des Signals unabhängig voneinander sind. Für den Fall von überlagerten Funksignalen entsteht eine ähnliche Situation: Es wird ein Signal über mehrere Pfade zum Empfänger geschickt. Die Phasen sind dabei ebenfalls gleichverteilt, da keine Synchronisation stattfindet. Es entsteht jedoch ein Leistungsgewinn durch die Summation, da ja nicht wie beim Rayleighkanal nur eine Quelle über mehrere Pfade zu einer Senke sendet, sondern mehrere Quellen über mehrere Pfade zu einer Senke senden. In Tabelle 3.1 sind die Terme aufgelistet, in denen sich das Rayleighmodell und nicht-kohärent überlagerte Funksignale entsprechen.

Die Rayleightheorie gibt die Statistik für die empfangenen Signale als komplexe Gaußverteilung an. Für den Fall der überlagerten Signale wird die Unab-

Rayleigh-Fading-Kanal	nicht-kohärente Überlagerung
viele Pfade	viele Sender
gleichförmig verteilte Phase wegen Mehrwegeausbreitung	gleichverteilte Phase wegen fehlender Synchronisation
insgesamt kein Leistungsgewinn	Leistungsgewinn durch Summation

Tabelle 3.1. Äquivalente Ausdrücke in der Theorie zum Rayleigh-Kanal und nicht-kohärenter Überlagerung

hängigkeit der N empfangenen Teilbeiträge vorausgesetzt und so kann man Mittelwert und Varianzen der summierten Signalkomponenten auf I-Achse ($\Re\{r_I(t)\}$) und der Q-Achse ($\Im\{r_I(t)\}$) auch summieren. Nimmt man weiter gleiche Empfangsleistung und Flat-Fading mit α an, so ergibt sich für die Signalmomente im Empfänger:

$$\begin{pmatrix} \mu_{\Im\{r(t)\}} \\ \mu_{\Re\{r(t)\}} \end{pmatrix} = N \cdot \underline{\mu}_{r_I(t)} = \underline{0}$$

$$\sigma_{\Im\{r(t)\}}^2 = \sigma_{\Re\{r(t)\}}^2 = N \cdot \sigma_{\Re\{r_I(t)\}}^2 = N\alpha^2 \sigma_{S_i}^2 \quad (3.7)$$

Hierbei wird mit $\sigma_{S_i}^2$ die Sendeleistung auf einer komplexen Signalkomponente verstanden

$$\sigma_{S_i}^2 := \frac{1}{2} |\underline{S}|^2 \quad (3.8)$$

Die Rayleigh-Theorie nimmt eine unendliche Zahl von Pfaden zwischen Empfänger und Sender an. In dem analogen Fall für überlagerte Signale korrespondiert dieser Parameter mit der Zahl der Sender. Diese Zahl wird nicht unendlich sein! Deshalb muss die Gültigkeit der Analogie auch für kleinere Anzahlen (also nicht unendlich viele Sender) überprüft werden. In Abbildung 3.5 sieht man die Dichteverteilungen der Q- und I- Empfangskomponenten für verschiedene Anzahlen von Sendern $N = 1, 2, 8$. Für $N = 1$ kann man die maximale Amplitude ± 1 erkennen. Für $N \rightarrow \infty$ sagt der zentrale Grenzwertsatz voraus, dass die Verteilung gaußisch sein wird; und zwar unabhängig von der ursprünglichen Verteilung für

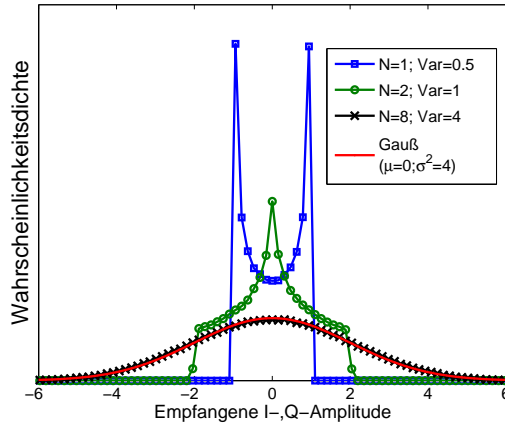


Abbildung 3.5. Die tatsächliche Verteilung für $N = 8$ ist der Gaußverteilung sehr ähnlich

$N = 1$. In Abbildung 3.5 kann man sehen, wie sich die Dichtefunktion für steigende N der Gaußverteilung nähert. Für $N = 8$ kann man den Unterschied kaum noch ausmachen. Deshalb wird im weiteren Verlauf die Rayleigh-Annahme auch für kleinere $n \leq \infty$ als gültig angenommen. Ein verbreitetes Maß, um Ähnlichkeiten zwischen Verteilungen anzugeben, ist der *Kullback-Leibler-Abstand*:

$$KL = \int p(x) \log \frac{p(x)}{q(x)} dx$$

In Abbildung 3.6 ist auch quantitativ zu sehen, wie sich die tatsächliche Verteilung der Gaußverteilung oder der Rayleigh-Verteilung über die Amplitude nähert.

3.1.2 Statistik im Empfänger

Die Äquivalenz zwischen der Theorie des Rayleighkanals und überlagerter Funk-signale aus mehreren, nicht-kohärenten Quellen wurde in dem vorangegangenen Abschnitt diskutiert und in Tabelle 3.1 veranschaulicht. Aufgrund des zentralen Grenzwertsatzes kann man annehmen, dass, wenn N Quellen ($N \geq 8$) das

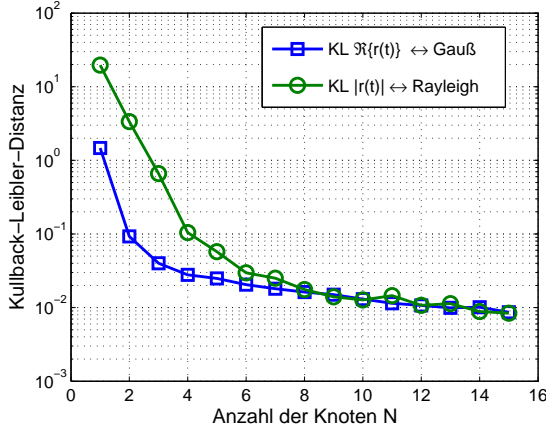


Abbildung 3.6. Der Kullback-Leibler-Abstand zwischen der tatsächlichen Verteilung auf I-Achse und der Gaußverteilung bzw. zwischen der tatsächlichen Amplitudenverteilung und der Rayleighverteilung

gleiche Signal S_i übertragen, das Empfangssignal einer komplexen Gaußstatistik entspricht $\sim \mathcal{N}(0, N\alpha^2\sigma_{S_i}^2\mathbf{I}^2)$. Den Rauschprozess kann man ebenfalls als unabhängig für die Signalkomponenten annehmen und bekommt hier ebenfalls eine komplexe Gaußverteilung: $\sim \mathcal{N}(0, \sigma_N^2\mathbf{I}^2)$. Das gesamte, empfangene Signal auf In-Phase und Quadraturkomponente entsteht in der Rayleightheorie dann durch die Faltung als Gaußverteilung:

$$f_{S_i}(\Re\{r(t)\} = x) = \frac{1}{\sqrt{2\pi(N\alpha^2\sigma_{S_i}^2 + \sigma_N^2)}} e^{-\frac{x^2}{2(N\alpha^2\sigma_{S_i}^2 + \sigma_N^2)}} \quad (3.9)$$

Betrachtet man nur die Amplitude u im Empfänger, also die Randdichte der Amplitude der komplexen Gaußverteilung, so entsteht eine Rayleighdichte:

$$f_{S_i}(|r(t)| = u) = \frac{u}{(N\alpha^2\sigma_{S_i}^2 + \sigma_N^2)} e^{-\frac{u^2}{2(N\alpha^2\sigma_{S_i}^2 + \sigma_N^2)}} \quad (3.10)$$

Zur vereinfachten Darstellung werden die Energien in einer neuen Variablen zusammengefasst:

$$N\alpha^2\sigma_{S_i}^2 + \sigma_N^2 =: \sigma_i^2 \quad (3.11)$$

und man hat mit (3.11) in (3.10) eine kompaktere Darstellung:

$$f_\sigma(u) = \frac{u}{\sigma^2} e^{-\frac{u^2}{2\sigma^2}} \quad (3.12)$$

3.1.3 Das Problem der Schwebungen

Bei inkohärenter Überlagerung kommt es im Empfänger zu Schwebungen. Dies ist in Abbildung 3.3 zu sehen. Diese Schwebungen entstehen durch unterschiedliche und sich zeitlich ändernde Phasenlagen der Teilbeiträge der verschiedenen Sender. In Abbildung 3.4 wurde dies ausführlich diskutiert. Für das System führen die Schwebungen zwar im Mittel nicht zu einer anderen Statistik, aber die Kurzzeitbetrachtung der überlagerten Signale kann deutlich anders aussehen. Dies hat vor allem für die praktische Verwendung extreme Auswirkungen. So könnten die Schwebungen gerade zu dem Zeitpunkt, in dem das Signal beobachtet wird, zu Auslöschungen führen. Im Einleitungsbeispiel in Kapitel 1.3 wäre diese Problematik sehr relevant, da die Beobachtungszeit der Knoten — und damit der Signale, die sie aussenden — durch das vorbeifliegende Flugzeug sehr kurz sein wird.

Exkurs ins Abtasttheorem

Die Schwebungen, die bei der Überlagerung von Signalen entstehen, sind kritisch für das Gesamtsystem. Diese Schwebungen sind allerdings bei der Überlagerung von Schwingungen unvermeidbar. In der Praxis bedeutet das, dass die Beobachtungszeit des resultierenden Signals mindestens so lang sein muss wie die Periode der Schwebung. Deshalb soll nun kurz die Periodenlänge eines solchen überlagerten Signals berechnet werden: Gegeben sind N Schwingungen $s_l(t) = \sin(\omega_l t + \varphi_l)$. Diese werden von den Quellen ausgesandt. Am Empfänger werden sie überlagert, also summiert:

$$s_\Sigma(t) = \sum s_l(t) \quad (3.13)$$

Nun wird die Periode T des summierten Signals bestimmt:

$$\begin{aligned} s_{\Sigma}(t) &= s_{\Sigma}(t+T) \\ \Leftrightarrow \sum s_l(t) &= \sum s_l(t+T) \\ \Leftrightarrow \sum \sin(\omega_l t + \varphi_l) &= \sum \sin(\omega_l t + \omega_l T + \varphi_l) \end{aligned} \quad (3.14)$$

Da die Schwingungen ω_l unterschiedlich sind (es findet ja keine Frequenzsynchronisation statt), kann (3.14) nur erfüllt werden, wenn die Einzelbeiträge übereinstimmen:

$$\begin{aligned} \forall l : \sin(\omega_l t + \varphi_l) &= \sin(\omega_l t + \omega_l T + \varphi_l) \\ \Leftrightarrow \omega_l t + \varphi_l &= \omega_l t + \omega_l T + \varphi_l - k2\pi \\ \Leftrightarrow 0 &= \omega_l T - k2\pi \\ \Leftrightarrow T &= \frac{k2\pi}{\omega_l} \\ \Leftrightarrow T_l &= \frac{2\pi}{\omega_l} \end{aligned} \quad (3.15)$$

Die kleinste Periode des Summensignals $s_{\Sigma}(t)$ wird erreicht als kleinstes gemeinsames Vielfaches (kgV) der Perioden der einzelnen Schwingungen:

$$T_{\Sigma} = k_l \cdot T_l; \quad k_l \in N \quad (3.16)$$

Hier genau entsteht die Problematik: Die Perioden der Einzelsignale können sich beliebig gering unterscheiden, somit kann die ganzzahlige Lösung von (3.16) zu einer beliebig großen Periode T_{Σ} für das Gesamtsignal führen. Je näher die Trägerfrequenzen der Quellen zueinander liegen, desto grösser wird die Periode des überlagerten Signals $s_{\Sigma}(t)$. Damit wird auch die nötige Beobachtungszeit eines Signals beliebig lange. Es wäre also hier von Vorteil, wenn die Trägerfrequenzen der Quellen weit auseinander liegen würden.

Bandspreizung mit Rauschsignatur

Zunächst stellt sich die Frage, was passiert, wenn man das Abtasttheorem verletzt, und das Summensignal $s_{\Sigma}(t)$ nicht mit der nötigen Periode beobachtet. In Abbildung 3.7 sind Empfangsleistungsverteilungen zu sehen, die zwei überlagerte Signale der Leistung 1 unter verschiedenen Bedingungen erzeugen. Man sieht,

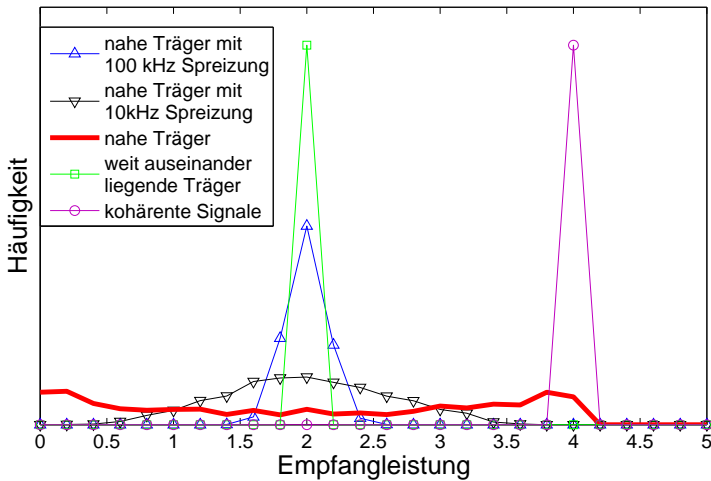


Abbildung 3.7. Verteilungen der Empfangleistung zweier überlagerter Signale: Bei Verletzung des Abtasttheorems ist die Verteilung nicht mehr zur Detektion geeignet.

dass die Leistungsverteilung bei Beobachtung eines Zeitausschnitts, der kürzer als die Periode ist, ihren höchsten Wert nicht auf dem Leistungswert 2 hat. Somit kann durch eine solche Überlagerung und Detektion kein sinnvolles System aufgebaut werden, welches die Leistungsaddition ausnutzt.

Zur Lösung dieses Problems gibt es nun verschiedene Ansätze. Die Sensorknoten könnten z.B. untereinander ihre Trägerfrequenzen messen und so einregeln, dass sie die geringst mögliche Gesamtperiode erzeugen. Das würde aber wiederum Frequenzsynchronisationen erfordern, die in dieser Arbeit nicht vorausgesetzt werden sollen. Ein anderer Lösungsvorschlag ist Bandspreizen mit Rauschsignalen. Hierbei wird den ausgesandten Signalen eine Rauschsignatur aufmoduliert. Diese Rauschsignatur ist ein möglichst breitbandiges, zufälliges Zeitsignal. Durch die zusätzliche Anwendung eines solchen Rauschsignals wird aus dem ausgesendeten Trägersignal ein ganzer Bandpassbereich, und der Effekt der Schwebungen aus überlagerten Trägern verringert sich. Durch die Bandspreizung wird erreicht, dass genug Frequenzanteile eines Signals weit genug von den Frequenzanteilen der anderen Quellen liegen. Somit wird die Periode der Summe (siehe (3.16)) verringert. Deshalb sollte beim Systemdesign die Bandspreizung so groß wie möglich sein. Der Idealfall wäre die Bandspreizung über das gesamte Spektrum, und die Überlagerung könnte beliebig kurz sein. Dies ist der theoretische Extremfall von Ultrabreitbandsignalen. Mit der Anwendung von Rauschsignalen als Signatur kann man quasi einen fließenden Übergang zwischen nicht-kohärenter Überlagerung von einfachen Schwingungen zu Breitbandsignalen erzeugen. Je größer die Bandbreite der Signatur ist, desto schmaler sind die Verteilungen der Empfangsleistungen um die ideale Leistungssumme verteilt. Je schmaler die Signatur ist, desto länger muss die Symboldauer sein, um sicher zu stellen, dass im Empfänger während der Dauer eines Symbols eine Periode des Summensignals beobachtet wurde. Symboldauer und Spreizbreite müssen somit in einem gemeinsamen Kompromiss festgelegt werden.

3.2 Verteiltes Senden

Dieses Szenario entsteht, wenn eine Gruppe von Sensorknoten Signale überlagert und diese Signale innerhalb der gleichen Gruppe wieder aufgenommen werden. Die Dämpfungen der Beiträge der einzelnen Sender α_i kann nun nicht mehr als identisch angenommen werden. Die Signalstatistiken auf der Amplitude ergeben sich analog zu (3.12):

$$f_{\sigma}(u) = \frac{u}{\sigma^2} e^{-\frac{u^2}{2\sigma^2}} \quad (3.17)$$

Mit den unterschiedlichen Kanaldämpfungen α_i wird deshalb σ_i folgendermaßen aufgelöst:

$$\sigma_i^2 := \sum_{l=1}^N \alpha_l^2 \sigma_{S_l}^2 + \sigma_N^2 \quad (3.18)$$

3.3 Gemeinsames Modell für ressourcenarme Knoten

Die Signalstatistiken im Empfänger (3.12) und (3.17) sind beide identisch. Lediglich die Zusammensetzung der Energiebeträge variiert wegen der unterschiedlichen Modelle für die Kanaldämpfung. In Abbildung 3.8 sind die beiden unterschiedlichen Szenarien und ihre Herleitung noch einmal zusammengefasst. Basie-

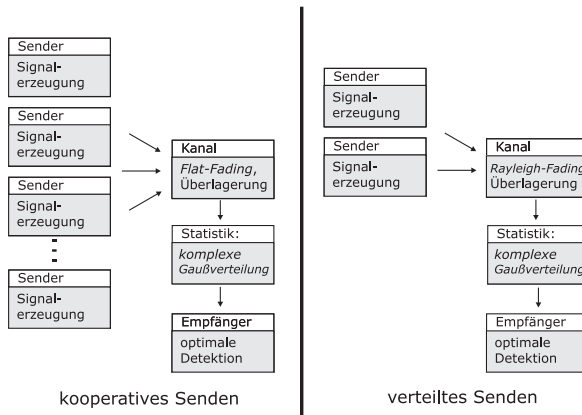


Abbildung 3.8. Unterschiedliche Systemmodelle für kooperatives und verteiltes Senden

rend auf der Rayleighstatistik im Empfänger soll nun eine optimale Übertragung hergeleitet werden. Die Übertragung spaltet sich in Signalformung (also Modulationsverfahren) und Detektion auf. Für die Detektion geht es dabei um die Detektion der Energie. Dies ist ein bekanntes Verfahren in der Signaltheorie, das im

folgenden Kapitel kurz dargestellt wird. Für die Signalformung wird eine umfassende Herleitung präsentiert, die in einem neuen Modulationsverfahren mündet. Dieses Modulationsverfahren positioniert die Symbole des mehrstufigen Übertragungsverfahrens derart, dass die Detektion bei Überlagerung optimal arbeiten kann.

3.4 Detektion des Rayleighparameters

Als einzige Größe im Empfänger bleibt die Leistung des empfangenen Signals. Deshalb wird im Empfänger eine einfache Maximum-Likelihood-Energiedetektion durchgeführt. Ziel der Detektion ist also eine Schätzung des Parameters σ des empfangenen Signals. Als Basis für den Schätzer werden mehrere Messwerte u_k der Amplitude verwendet. Setzt man voraus, dass diese unabhängig voneinander sind, so kann man den ML-Schätzer herleiten. Ausgangspunkt ist die Dichtefunktion im Empfänger (3.12):

$$f_{\sigma}(u) = \frac{u}{\sigma^2} e^{-\frac{u^2}{2\sigma^2}}$$

der optimale ML-Schätzer ist:

$$\arg \max_{\sigma} f_{\sigma}(u) \quad (3.19)$$

Die Optimalität wird also erreicht, wenn gilt:

$$\frac{\partial f_{\sigma}(u)}{\partial \sigma} = 0 \quad (3.20)$$

Mit unabhängigen Abtastwerten $u_k; k = 1..N$, die im Vektor \mathbf{u} zusammengefasst werden, ist die Verbunddichte:

$$f_{\mathbf{U}}(\mathbf{u}) = \prod_{i=1}^N f_{\sigma}(u_i) = \prod_{i=1}^N \frac{u_i}{\sigma^2} e^{-\frac{u_i^2}{2\sigma^2}} = \frac{\prod_{i=1}^N u_i}{\sigma^{2N}} \cdot e^{-\frac{1}{2\sigma^2} \sum_{i=1}^N u_i^2} \quad (3.21)$$

Für die Likelihoodschätzung ist es ein bekanntes Hilfsmittel, eine Logarithmierung durchzuführen. Der Logarithmus ist stetig und streng monoton steigend, deshalb kann man (3.20) erweitern und bekommt so eine neue Vorschrift:

$$\frac{\partial (\ln f_{\sigma}(u))}{\partial \sigma} = 0 \quad (3.22)$$

Die Anwendung des Logarithmus auf die Rayleigh-Verbunddichte führt zu:

$$\begin{aligned} \ln(f_{\mathbf{U}}(\mathbf{u})) &= \ln\left(\frac{\prod_{i=1}^N u_i}{\sigma^{2N}}\right) + \ln e^{-\frac{1}{2\sigma^2} \sum_{i=1}^N u_i^2} \\ &= \sum_{i=1}^N \ln(u_i) - \ln(\sigma^{2N}) - \frac{1}{2\sigma^2} \sum_{i=1}^N u_i^2 \end{aligned} \quad (3.23)$$

Ableitung von (3.23) nach σ führt zu:

$$\begin{aligned} \frac{\partial(\ln f_{\sigma}(u))}{\partial \sigma} &= 0 - \frac{2N}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^N u_i^2 \stackrel{!}{=} 0 \\ \Leftrightarrow \tilde{\sigma}^2 &= \frac{1}{2N} \sum_{i=1}^N u_i^2 \end{aligned} \quad (3.24)$$

Mit (3.24) ist die ML-Schätzvorschrift für den Rayleighparameter hergeleitet. Im einfachsten Fall werden nicht N , sondern nur ein einziger Messwert verwendet, dann kontrahiert (3.24) zu:

$$\tilde{\sigma}^2 = \frac{u^2}{2}$$

Mit diesen Schätzvorschriften kann der Empfänger optimal arbeiten und verwendet dabei — wie gefordert — nur die empfangene Amplitude.

3.5 ESK

ESK steht für „energy shift keying“. Der Name für dieses Modulationsverfahren reflektiert eine seiner wichtigsten Eigenschaften: Es basiert ausschließlich auf der Verwendung von Energiewerten (oder Amplitudenwerten). Dies ist nicht zu verwechseln mit ASK (amplitude shift keying), dass zwar auch auf der Verwendung von Amplitudenwerten basiert, aber nicht wie ESK rotationsinvariant im komplexen Signalraum ist. Die Frage, die das nun herzuleitende Modulationsverfahren beantworten soll, ist, wie die Signalkonstellation zu wählen ist, damit bei Überlagerung von verschiedenen Quellen und der Detektion mit dem ML-Schätzer aus

Abschnitt 3.4 eine optimale Detektion einer mehrstufigen Modulation erreicht werden kann.

Da der einzige freie Parameter für die mehrstufige Modulation die Leistung σ_i^2 der ausgesendeten Signale S_i ist, müssen also diese Leistungen im Signalraum so platziert werden, dass eine bestmögliche Erkennung im Empfänger entsteht. Im Empfänger wird die Entscheidung für die Detektion des Symbols S_i als Hypothese H_i bezeichnet. Als Optimalitätskriterium wird gewählt, dass die richtige Erkennung der Signale maximal werden soll:

$$P = \sum_{i=0}^{M-1} P(S_i)P(H_i|S_i) \rightarrow \max \quad (3.25)$$

Um die optimale Signalkonstellation (oder besser Energiekonstellation) herleiten zu können, müssen zunächst die Detektionsgrenzen für den Empfänger festgelegt werden. Diese Grenzen legen fest, für welches Symbol sich der Empfänger entscheidet (H_i), wenn er eine gewisse Sendeleistung σ^2 empfängt. Die Entscheidungsgrenzen werden als Maximum-Likelihood-Entscheidung festgelegt.

Behauptung: Die optimalen Entscheidungsgrenzen für den ML-Empfänger liegen in den Schnittpunkten der Rayleighdichten für die verschiedenen Rayleighparameter (oder Empfangsenergien) σ_i

Beweis: Für diesen Beweis wird eine graphische Illustration herangezogen. In Abbildung 3.9 sieht man typische Rayleighverteilungen für verschiedene Parameter (oder Energien) σ_i mit $\sigma_1^2 < \sigma_2^2 < \sigma_3^2$. Zunächst soll es nun um die Entscheidungsschwelle γ_{12} gehen, die die Hypothesen H_1 (=Entscheidung für S_1) und H_2 trennt. Diese Entscheidungsschwelle wird am Schnittpunkt der zugehörigen Rayleighdichten angesetzt. Bewegt man nun diese Entscheidungsgrenze nach links oder rechts, so wird die gesamte Fehlerwahrscheinlichkeit in beiden Richtungen erhöht. Die Fehleranteile, die durch die Entscheidungsgrenze γ_{12} beeinflusst werden, können in zwei Gruppen unterteilt werden:

$$P_{e,S_1} = P(H_1|S_2) + P(H_1|S_3)$$

$$P_{e,S_2} = P(H_2|S_1) + P(H_2|S_3)$$

$$P(\text{Veränderung des Gesamtfehlers durch diese Grenze}) =$$

$$P_{e,S_1} + P_{e,S_2}$$

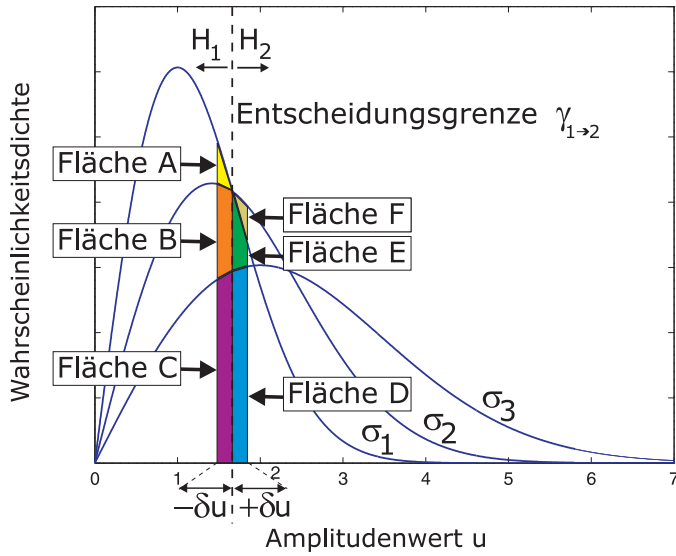


Abbildung 3.9. Optimale Entscheidungsgrenze $\gamma_{1 \rightarrow 2}$

Wenn nun die Schwelle ∂u nach links bewegt wird, so gibt es einige Änderungen in den Fehleranteilen:

$$\begin{aligned} P_{e,S_1,-\partial x} &= P_{e,S_1} - (B+C) - (C) \\ + P_{e,S_2,-\partial x} &= P_{e,S_2} + (A+B+C) + (C) \end{aligned}$$

$$P_{e,S_1,-\partial x} + P_{e,S_2,-\partial x} = P_{e,S_1} + P_{e,S_2} + (A) \quad (3.26)$$

wobei die Flächen $(A) - (F)$ als Integration über die Wahrscheinlichkeitsdichte der Empfangsamplitude entstehen. Wenn die Schwelle ∂u nach rechts bewegt wird, gibt es auch einige Veränderungen:

$$\begin{aligned} P_{e,S_1,+\partial x} &= P_{e,S_1} + (D+E+F) + (D) \\ + P_{e,S_2,+\partial x} &= P_{e,S_2} - (D+E) - (D) \end{aligned}$$

$$P_{e,S_1,+\partial x} + P_{e,S_2,+\partial x} = P_{e,S_1} + P_{e,S_2} + (F) \quad (3.27)$$

In beiden Fällen wird der Gesamtfehler größer, wenn die Schwelle von dem Schnittpunkt der Dichten wegbewegt wird. Das ist das typische Maximum-Likelihood-Verhalten. Die Entscheidungsgrenzen im Empfänger können also in den Schnittpunkten der Rayleighdichten für die verschiedenen Empfangsleistungen festgelegt werden. Mit der Rayleighdichte (3.12) und zwei benachbarten Leistungen σ_i und σ_{i+1} erhält man die Entscheidungsschwelle

$$\begin{aligned} \frac{u}{\sigma_i^2} e^{-\frac{u^2}{2\sigma_i^2}} &= \frac{u}{\sigma_{i+1}^2} e^{-\frac{u^2}{2\sigma_{i+1}^2}} \\ \Leftrightarrow u^2 &= \ln\left(\frac{\sigma_i^2}{\sigma_{i+1}^2}\right) \cdot \frac{2\sigma_i^2\sigma_{i+1}^2}{\sigma_i^2 - \sigma_{i+1}^2} =: \gamma_{i \rightarrow i+1}^2 \end{aligned} \quad (3.28)$$

Mit Hilfe dieser Schwelle kann nun die Entdeckungswahrscheinlichkeit maximiert werden:

$$\begin{aligned} P &= \sum_{i=0}^{M-1} P(S_i) \cdot P(H_i|S_i) \quad |P(S_i) = 1/M \\ P &= \frac{1}{M} \sum_{i=0}^{M-1} \int_{\gamma_{i-1 \rightarrow i}}^{\gamma_{i \rightarrow i+1}} f_{\sigma_i^2}(u) du \\ \text{mit } f_{\sigma_i^2}(u) &= \frac{u}{\sigma_i^2} e^{-\frac{u^2}{2\sigma_i^2}} \end{aligned} \quad (3.29)$$

Mit (3.28) und (3.12) in (3.29):

$$\begin{aligned}
 M \cdot P = P_d &= 1 - e^{\frac{\ln(\frac{\sigma_1^2}{\sigma_0^2}) \cdot \sigma_1^2}{\sigma_0^2 - \sigma_1^2}} + \\
 &+ \sum_{i=0}^{M-3} \left[e^{\frac{\ln(\frac{\sigma_{i+1}^2}{\sigma_i^2}) \cdot \sigma_i^2}{\sigma_i^2 - \sigma_{i+1}^2}} - e^{\frac{\ln(\frac{\sigma_{i+2}^2}{\sigma_{i+1}^2}) \cdot \sigma_{i+2}^2}{\sigma_{i+1}^2 - \sigma_{i+2}^2}} \right] + \\
 &+ e^{\frac{\ln(\frac{\sigma_{M-1}^2}{\sigma_{M-2}^2}) \cdot \sigma_{M-2}^2}{\sigma_{M-2}^2 - \sigma_{M-1}^2}}
 \end{aligned} \tag{3.30}$$

Es gilt nun, die Leistungen σ_i^2 so zu wählen, dass der Ausdruck (3.30) maximal wird; dabei sind σ_0^2 und σ_{M-1}^2 die minimalen und maximalen Empfangsenergien und werden durch die Energiebeschränkungen festgelegt. Sie sind deshalb von dem Optimierungsprozess ausgeschlossen. Mit der Festlegung $\sigma_0^2 = 0$ und der *virtuellen* Leistung $\sigma_M^2 = \infty$ kann der Ausdruck (3.30) kompakter geschrieben werden:

$$P_d = \sum_{i=0}^{M-2} \left[e^{\frac{\ln(\frac{\sigma_{i+1}^2}{\sigma_i^2}) \cdot \sigma_i^2}{\sigma_i^2 - \sigma_{i+1}^2}} - e^{\frac{\ln(\frac{\sigma_{i+2}^2}{\sigma_{i+1}^2}) \cdot \sigma_{i+2}^2}{\sigma_{i+1}^2 - \sigma_{i+2}^2}} \right] \tag{3.31}$$

Um die Optimierung durchzuführen, wird der Summenausdruck (3.31) in Teilschritten zerlegt. Es werden zunächst alle Teilbeiträge herausgelöst, die σ_i^2 enthalten. In diesem Ausdruck werden σ_{i-1}^2 und σ_{i+1}^2 als Parameter verstanden:

$$\begin{aligned}
 P_{d, \sigma_{i-1}^2, \sigma_{i+1}^2}(\sigma_i^2) &= -e^{\frac{\ln(\frac{\sigma_i^2}{\sigma_{i-1}^2}) \cdot \sigma_i^2}{\sigma_{i-1}^2 - \sigma_i^2}} + e^{\frac{\ln(\frac{\sigma_i^2}{\sigma_{i-1}^2}) \cdot \sigma_{i-1}^2}{\sigma_{i-1}^2 - \sigma_i^2}} \\
 &- e^{\frac{\ln(\frac{\sigma_i^2}{\sigma_{i+1}^2}) \cdot \sigma_i^2}{\sigma_i^2 - \sigma_{i+1}^2}} + e^{\frac{\ln(\frac{\sigma_i^2}{\sigma_{i+1}^2}) \cdot \sigma_{i+1}^2}{\sigma_i^2 - \sigma_{i+1}^2}}
 \end{aligned} \tag{3.32}$$

Da die optimale Wahl von σ_i nur von den Nachbarn σ_{i-1} und σ_{i+1} abhängt, kann man das lokale Maximum von σ_i in Abhängigkeit seiner Nachbarn σ_{i-1}

und σ_{i+1} bestimmen. Wenn die optimale Wahl von σ_i in Bezug auf σ_{i-1} und σ_{i+1} gefunden ist, so kann man anschließend den Index von i auf $i+1$ verschieben und mit dem gleichen Ergebnis die Wahl von σ_{i+1} optimieren. Da es keine weiteren Abhängigkeiten gibt, kann man durch wiederholte Anwendung dieses Prozesses die gesamte Summe iterieren und erzeugt so das globale Maximum. In Anhang A wird gezeigt, dass der Ausdruck (3.32) optimiert wird, wenn die σ_i durch folgende Bedingung miteinander verknüpft sind:

$$\sigma_i^2 = k \cdot \sigma_{i-1}^2 \quad \text{mit } 1 < i < M ; \quad k > 1, k \in \mathbb{R} \quad (3.33)$$

Für die optimale Signalkonstellation muss nun nur noch (3.11) in (3.33) eingesetzt werden, und nach der Sendeleistung aufgelöst werden:

$$\sigma_{S_i}^2 = \frac{k(\sigma_N^2 + N\alpha^2 \sigma_{S_{i-1}}^2) - \sigma_N^2}{N\alpha^2} \quad (3.34)$$

Für diese Lösung ist es allerdings notwendig, dass die Sender die Anzahl der anwesenden Geräte N und den Rauschlevel σ_N^2 im Empfänger kennen. Dies ist sicherlich nicht immer der Fall. Aber man kann (3.34) vereinfachen, wenn man annimmt, dass $\sigma_{S_{M-1}}^2 \gg \sigma_{S_0}^2 \gg \sigma_N^2$:

$$\sigma_{S_i}^2 \approx k \cdot \sigma_{S_{i-1}}^2 \quad \text{und} \quad k \approx \sqrt{\frac{\sigma_{S_{M-1}}^2}{\sigma_{S_0}^2}}^{(M-1)} \quad (3.35)$$

Mit dieser Vorschrift ist eine einfache, praktische Näherung für die Signalkonstellation erreicht. Abbildung 3.10 zeigt beispielhaft Signalkonstellationen für ESK und ASK, PSK. Die Besonderheit von ESK ist, dass es rotationsinvariant ist. Die Signalpunkte in ESK legen nur die Wahl der Energie, also der Entfernung von Ursprung fest, nicht den Winkel in der komplexen Signalebene.

3.6 Energienormierung

Der ML-Detektor (3.24) kann die Leistung eines empfangenen Signals schätzen. Für ESK ist es allerdings relevant, die empfangene Leistung auch einem Symbol zuzuordnen. Deshalb muss die minimale σ_0^2 und die maximale σ_{M-1}^2 Empfangsleistung bekannt sein, damit die Entscheidungsgrenzen $\gamma_{i \rightarrow i+1}$ festgestellt werden

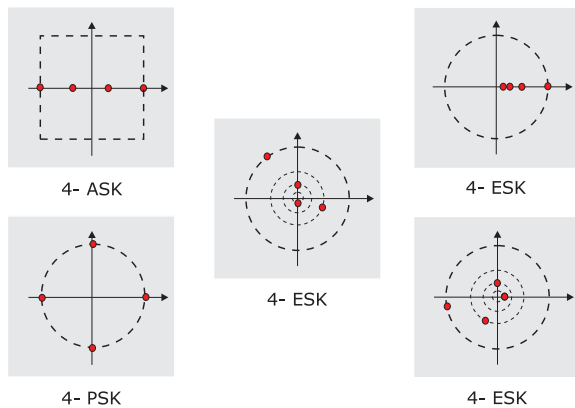


Abbildung 3.10. Signalkonstellation in der komplexen Signalebene: Vergleich traditioneller ASK und PSK mit ESK für überlagerte Funksignale. Die Signalkonstellation für ESK ist rotationsinvariant – die Signalepunkte können auf Kreisen beliebig platziert werden.

können. Die maximale und minimale Leistung kann z.B. durch eine Trainingsphase in der Präambel einer Nachricht geschätzt werden. Mit Hilfe von (3.35) und der bekannten Anzahl M der Symbole kann dann k bestimmt werden. Mit bekanntem k und der Rekursion (3.33) in (3.28) erhält man (ebenfalls rekursiv) die Entscheidungsschwellen (Entscheidungsenergien) im Empfänger:

$$\begin{aligned}\gamma_{i \rightarrow i+1}^2 &= \ln \left(\frac{\sigma_i^2}{k\sigma_i^2} \right) \cdot \frac{2\sigma_i^2 k \sigma_i^2}{\sigma_i^2 - k\sigma_i^2} \\ &= \ln \left(\frac{1}{k} \right) \cdot \frac{2k\sigma_i^2}{1-k}\end{aligned}\quad (3.36)$$

Mit der Bekannten maximalen Energie σ_{M-1}^2 können dann die Entscheidungsschwellen gefunden werden.

3.7 Zusammenfassung

In der Herleitung der Signalstatistiken wurden Analogien zur Theorie des Rayleigh-Fading-Kanals gefunden. Diese in Tabelle 3.1 zusammengestellte Analogie führt letztlich auf das Modell der Statistik der überlagerten Signale (3.10). Diese stellt nun die Basis für die Herleitung der optimalen Signalkonstellation dar. Mit (3.34) ist eine optimale Signalkonstellation für eine M-stufige Modulation für überlagerte Funksignale hergeleitet worden. Das Optimalitätskriterium während der Herleitung (3.25) für maximale Erkennung wurde begleitet mit der Einschränkung auf eine maximale Energie σ_{M-1}^2 . Weiterhin wird mit (3.35) eine praktische Annäherung gegeben und mit (3.36) die Vorschrift, wie die Empfänger die Detektionsgrenzen wählen sollten. Es ist ebenfalls wichtig daran zu erinnern, dass für die gesamte Herleitung lediglich die Zeitsynchronisation vorausgesetzt wird und keine komplexwertige Signalverarbeitung notwendig ist. Die Oszillationen, die durch die Inkohärenz der Quellen entstehen, und die Summation der vielen Signale verursachen die Rayleighstatistik der Empfangsamplitude, welche als Basis für die Herleitung der optimalen Konstellation dient. Wenn Signale im Empfänger nur über kurze Zeit beobachtet werden, kann es sein, dass sich die Verteilung solcher Stichproben deutlich von der eigentlichen Statistik (3.10) unterscheidet. In einem solchen Fall wäre die Detektion fehlerhaft. Um diesen Einflüssen, die für die hier gezeichnete Anwendung besonders z.B. durch Schwebungen entstehen, entgegenzuwirken, wurde in Kapitel 3.1.3 ein Bandspreizverfahren vorgeschlagen, das die Kurzzeitstatistik im Sinne der Detektion verbessert.

4 Implementierung

In den vorangegangenen Kapiteln wurden detaillierte Analysen und Herleitungen durchgeführt, die die Grundsteine für überlagerte Funksignale liefern. Die vorliegende Arbeit erhebt den Anspruch, dass alle dargestellten Zusammenhänge und Forschungsergebnisse tatsächlich praxisrelevant sind. Die Arbeit ist deshalb auch speziell auf die Anforderungen von leistungsschwachen Sensorknoten ausgerichtet. Um die Praxisrelevanz zu zeigen und die theoretischen Arbeiten zu untermauern, werden in dem nun folgenden Kapitel einige Implementierungen gezeigt. Diese Implementierungen wurden auf pPart Particle Computer durchgeführt. In der wissenschaftlichen Literatur zu überlagerten Funksignalen gibt es bisher keinen Hinweis auf eine Implementierung von Mechanismen, die von den verschiedenen Autoren vorgeschlagen wurden. Im folgenden Kapitel wird eine funktionierende Implementierung von überlagerten Funksignalen vorgestellt, die bisher in diesem Forschungsbereich in keiner veröffentlichten Arbeit behandelt wurde.

Die vorangegangenen Kapitel beschäftigten sich mit der Analyse und der theoretischen Herleitung von Modellen und Übertragungsmechanismen. In diesem Kapitel soll nur die Implementierbarkeit der herausgearbeiteten Vorschläge untersucht werden. Die Implementierung wird nicht als Simulation, sondern direkt auf einer Hardwareplattform vorgenommen werden, um zu zeigen, dass die erarbeiteten Beiträge tatsächlich implementierbar sind. Es gibt in der gesamten relevanten Literatur bisher keinen Hinweis darauf, dass jemals überlagerte Funksignale auf echter Hardware implementiert wurden. In der vorliegenden Arbeit wird neben theoretischen, notwendigen Untersuchungen und Simulationen, auch eine Implementierung auf der Zielpattform tatsächlich realisiert.

Es gilt nun folgende Aspekte der vorangegangenen Kapitel zu überprüfen:

1. Wie sehen im Bandpass überlagerte Funksignale tatsächlich aus?
2. Verhalten sich überlagerte Signale und Leistungen tatsächlich additiv?
3. Kann man mit überlagerten Funksignalen tatsächlich die in dem Motivationsbeispiel skizzierte Reichweitenproblematik lösen?

Die Implementierung und Untersuchung der Fragestellungen werden auf der Hardwareplattform pPart Particle Computer durchgeführt. Diese Plattform hat den Vorteil, dass sie sämtliche Protokollschichten von der Applikation bis zur Basisbandsignalisierung in Software implementiert. Somit ist sie geeignet, genaue Untersuchungen der Effekte von überlagerten Funksignalen zu analysieren und kann durch die Softwareimplementierung auch viele andere Hardwarekonstellationen simulieren.

4.1 Wie sehen im Bandpass überlagerte Funksignale tatsächlich aus?

Die erste Frage, die mit den Implementierungen beantwortet werden soll, ist weniger analytischer Natur, sondern soll ein lebendiges Bild der diskutierten Thematik erzeugen. Überlagerte Funksignale findet man noch nicht in Lehrbüchern, deshalb sollen möglichst viele Aspekte dieses neuen Forschungsbereiches präsentiert werden, um dem interessierten Leser ein möglichst vollständiges und anschauliches Bild zu liefern. Der erste Versuchsaufbau ist nun ziemlich einfach: Mit fünf Particle Computer pParts wurde ein Protokoll in Software implementiert, das Funksignale überlagert. Da alle Funkknoten in Reichweite jeweils aller anderen Partner sind, ist die Synchronisation und Organisation dieses Protokolls

trivial. Zur Visualisierung wurde an eines der pParts ein Oszilloskop angeschlossen. Auf dem Oszilloskop wurde so der Amplitudenwert im empfangenen Basisband darstellt. In Abbildung 4.1 sieht man eine Momentaufnahme von überlager-

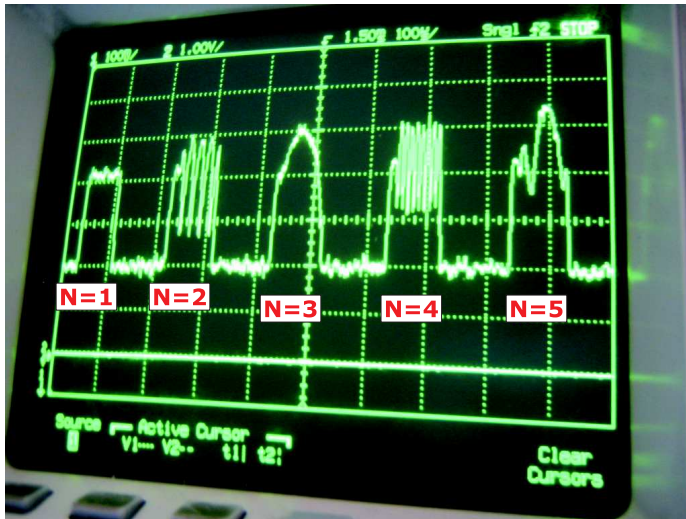


Abbildung 4.1. Überlagerung von Funksignalen

ten Funksignalen. Man sieht fünf aufeinanderfolgende Signale, die von links nach rechts eine einfache ($N=1$, also keine), zweifache ($N=2$), dreifache ($N=3$), vierfache ($N=4$) und fünffache ($N=5$) Überlagerung zeigen. Die ausgesandten Signale waren Rechtecksignale der Amplitude im Basisband, die von den Sendern auf 868.35 Mhz versendet wurden. Bei der zweifachen Überlagerung kann man deutlich eine Schwingung in der Amplitude im Empfänger erkennen. Dies sind genau die Schwebungen, wie sie in Kapitel 3.1.3 diskutiert wurden. Man kann prinzipiell sehen, dass die Gesamthöhe der angezeigten Signale mit jedem Dazukommen eines zusätzlichen Teilnehmers steigt. Durch die Schwebungen und die zeitlich veränderlichen konstruktiven und destruktiven Überlagerungen scheint es jedoch nicht trivial, die empfangene Leistung klar auf die Anzahl der Quellen zurückzuführen. Ob diese Summierung und damit der Überlagerungseffekt tatsächlich quantitativ zu messen ist, soll das nächste Kapitel untersuchen.

4.2 Verhalten sich überlagerte Signale und Leistungen tatsächlich additiv?

Für die Signalmodellierung im Kapitel ist es eine wichtige Voraussetzung, dass sich die Leistungen von überlagerten Funksignalen additiv verhalten. Diese Fragestellung wurde in einer Studienarbeit [21] bearbeitet. In den Versuchsaufbauten wurden auch pParts verwendet, die in verschiedenen Umgebungen und Entfernungen platziert wurden. Typische Entfernungen waren ein bis drei Meter. Die Versuche wurden im Labor und in Büros durchgeführt. Für die Charakteristik des verwendeten Transceivers¹ mussten zunächst einige Vorverarbeitungsschritte durchgeführt werden. Das Signal im Basisband wird von dem verwendeten Transceiver in logarithmischer Skala als Spannungswert ausgegeben. Zur Analyse dieses Signals muss es wieder auf eine lineare Skala abgebildet werden. In Abbildung 4.2 sieht man auf der linken Seite die Spannungsmesswerte des Transceivers für 1- bis 5-fache Überlagerung (von oben nach unten) und auf der rechten Seite die Werte auf eine lineare Skala gestreckt. Damit alle Teilnehmer mit gleicher Leistung am Empfänger erscheinen, wurde eine Leistungsnormierung als Protokoll implementiert. Wichtig ist hier zu bemerken, dass diese Normierung nicht für das eigentliche System notwendig ist. Für die mehrstufige ESK-Übertragung z.B. können über die Trainingsphase einer Präambel die nötigen Parameter bestimmt werden — die Entfernungen und Kanaldämpfungen wären für alle Quellen gleich. In dem Versuchsaufbau ist jedoch dieses Kanalmodell nicht gültig, und die Implementierung nutzt auch keine Normalisierung durch Präambeln, deshalb war die zusätzliche Energienormierung für die detaillierte Analyse in den Versuchen hilfreich. Zur Auswertung im Empfänger wurden nicht direkt die Abtastwerte des gestreckten Empfangssignals, sondern ein Energiemaß verwendet. Das Energiemaß wurde als Summe der Abtastwerte bestimmt. Zurückblickend auf die Anforderungen an die Zielform ist die Verwendung eines solchen Maßes sehr sinnvoll. Es kann durch einfache, analoge Aufbauten (z.B. einen Integrator) realisiert werden. In Abbildung 4.3 sind Aufnahmen von 5 Versuchen zu sehen. Die Entfernungen zum Empfänger waren in den Versuchen unterschiedlich, was zu unterschiedlichen, absoluten Energiewerten im Empfänger führt. Bei Versuch 1 sind die Entfernungen zwischen Sendern und Empfänger z.B. wesentlich größer als bei Versuch 5, was zu insgesamt niedrigeren Werten in Versuch 1 führt. Bei den Versuchen wurden jeweils bis zu 5 Signale überlagert und im Empfänger verarbeitet. Die Energiemaße steigen mit zunehmender Anzahl von Sendern

¹RFM TR1001

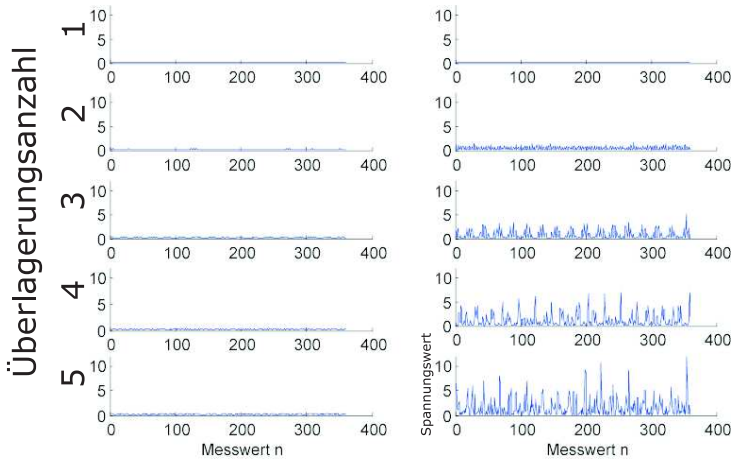


Abbildung 4.2. Transformation der Messwerte von einer logarithmischen (links) in eine lineare (rechts) Skala; aus [21, Abbildung 18]

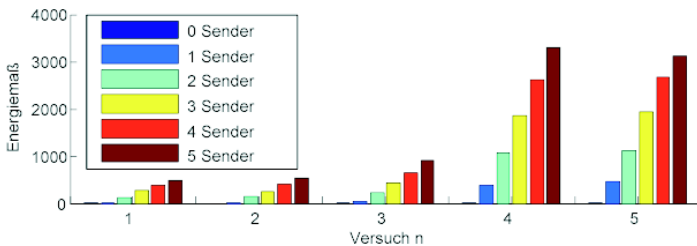


Abbildung 4.3. Ermittelte Energiemaße bei verschiedenen Versuchen; aus [21, Abbildung 20]. Die Versuche sind aufsteigend sortiert nach der mittleren empfangenen Energie. Diese Energie hängt z.B. von der geometrischen Position der Knoten ab, die in den Versuchen unterschiedlich war.

4.3 Kann man mit überlagerten Funksignalen das Sensor-Reach-back-Problem lösen?

deutlich. Diese Feststellung gilt für alle Versuche. Die Energiemaße wurden dann durch einen Schätzer (Schätzer für den Rayleighparameter: (3.24)) und eine Klassifikation auf die Anzahlen der Sender zugeordnet. In über 2000 Einzelversuchen wurde diese Klassifikation praktisch evaluiert. In Abbildung 4.4 ist das Ergebnis der Klassifikationen zu sehen. Der Gesamtfehler durch Fehlklassifikation beträgt 15%, wobei Schätzungen mit einem Fehler größer als 1 nur in 1.5% aller Fälle auftraten. Der relative große Gesamtfehler erklärt sich durch vielfältige Einflüsse

		Geschätzte Anzahl					
tatsächliche Anzahl		0	1	2	3	4	5
	0	287	0	0	0	0	0
	1	0	327	3	0	0	0
	2	0	0	330	0	0	0
	3	0	0	32	321	14	19
	4	0	0	11	69	211	124
	5	0	0	0	17	39	220

Abbildung 4.4. Confusion-Matrix für die Erkennung von mehrfach überlagerten Funksignalen

aus der Hardware und Software, besonders auch durch die verwendete Leistungsregelung in der Normierung, die aufgrund der Einschränkungen der Hardware nur in sehr groben Stufen möglich war. In dieser ersten Implementierung war das Ziel nicht, die Detektion und Verarbeitung ideal und genau zu realisieren, sondern prinzipiell zu zeigen, dass eine Implementierung auf echter Hardware möglich ist. Im Laufe der Anwendungskapitel dieser Arbeit wird noch eine weitere Implementierung von überlagerten Funksignalen vorgestellt, die anwendungsbezogener ist und fast ideale Ergebnisse zeigt.

4.3 Kann man mit überlagerten Funksignalen das Sensor-Reach-back-Problem lösen?

Dieser Fragestellung wurde ebenfalls mit einem Versuchsaufbau nachgegangen. Er ist in Abbildung 4.5 zu sehen. Auf einem Sendepult sind vier pPart Sensorkno-

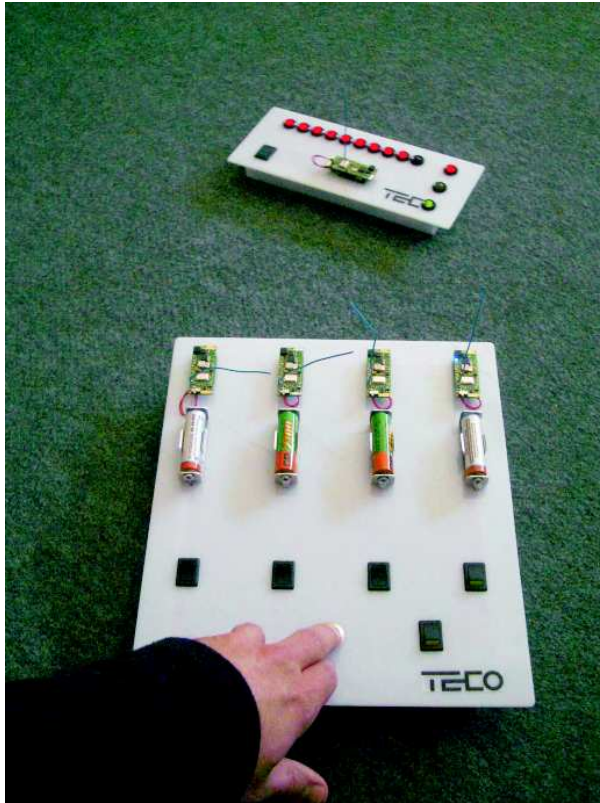


Abbildung 4.5. Versuchsaufbau zu Sensor-Reach-Back-Problem; aus [41]

ten aufgebaut. Diese können mit Schaltern eingeschaltet werden. Der Empfänger zeigt die empfangene Leistung an. Schaltet man mehr Sender ein, so erhöht sich die Empfangsleistung im Empfänger. Wenn eine gewisse Schwelle im Empfänger überschritten ist, so löst der Empfänger einen Alarm aus. Mit diesem Versuchsaufbau wurde auf zwei Konferenzen (PERVASIVE 2006, [41] und INSS 2006 [42]) gezeigt, dass es möglich ist, mit überlagerten Funksignalen die Reichweite von Sensorknoten zu erhöhen.

4.4 Zusammenfassung

In diesen ersten Implementierungen von überlagerten Funksignalen auf Sensorknoten wurde gezeigt, dass die Theorien und Modelle der vorangegangenen Kapitel tatsächlich nutzbare Ergebnisse auf realer Hardware erzeugen. Nur mit den starken Einschränkungen für die Zielplattform und den sich daraus ergebenden einfachen Modellen, konnte diese Implementierung auf einer realen Sensornetzplattform gelingen. Sämtliche Vorschläge anderer Autoren, die in Kapitel 2.2 vorgestellt wurden, wären an dieser Implementierungshürde gescheitert.

Als Implementierungsbasis wurde die pPart Particle Computer Plattform gewählt, die als typische Zielplattform verstanden werden kann. Bei der Signalerzeugung wurde (wie in Kapitel 3.1.3 beschrieben) eine Rauschsignatur eingesetzt. In Kapitel 3.1.3 wurde gezeigt, dass sich mit zunehmender Bandbreite der Signatur die Detektion verbessert. Die maximale Bandbreite in den Versuchen war automatisch durch die Bandbreite des Transceivers auf 125kHz festgelegt. Zusätzlich zu dieser Signalformung wurde eine Leistungsnormierung durchgeführt, die alle Empfangssignale auf etwa gleiche Leistungen eingeregelt hat. Damit sollten die Kanalmodelle (Flat-Fading und Rayleighkanal) aus Kapitel 3 dargestellt werden, in denen alle Signale mit etwa gleicher Leistung im Empfänger eintreffen. Für die Detektion wurde zunächst eine Signalaufbereitungsstufe eingesetzt, die die Ausgabespannungen des Transceivers wieder in Empfangsleistungen überführt. Anschließend wurde das Ergebnis in einer ML-Klassifikation bewertet, wie es in Kapitel 3.4 vorgeschlagen wurde. In den Versuchen wurde 85% richtige Klassifikation bei Überlagerung von bis zu 5 Signalen erreicht. Mit diesen Versuchen wurde praktisch gezeigt, dass sich die Leistungen überlagerter Funksignale additiv verhalten, und somit die Signalmodelle in Kapitel 3 brauchbar sind.

5 AwareCon: Einsatz von überlagerten Funksignalen zur Verbesserung von Netzwerkperformance

Das Funkprotokoll AwareCon für drahtlose Sensornetze wurde im Zusammenhang mit der vorliegenden Arbeit am Telecooperation Office der Universität Karlsruhe entwickelt. Es wurde speziell für Anwendungen im Bereich Ubiquitous Computing entworfen und trägt deshalb Eigenschaften, die den typischen Anforderungen gerecht werden. Es ist z.B. speziell auf Mobilität von Netzwerkteilnehmern zugeschnitten. Die zwei charakterisierenden Mechanismen von AwareCon sind Zeitsynchronisation und Kanalzugriff und basieren auf überlagerten Funksignalen. Diese Zeitsynchronisation erreicht eine Uhrensynchronität von mindestens $4\mu\text{s}$ zwischen allen Netzwerkteilnehmern. Das Kanalzugriffsverfahren ist eine Adaption eines bekannten Verfahrens aus der verdrahteten Welt, das mit AwareCon zum ersten Mal Einzug in die drahtlose Welt nimmt. Mit diesem Zugriffsverfahren können hohe Geräteanzahlen verarbeitet und Echtzeitanforderungen erfüllt werden. AwareCon ist als Funkprotokoll für Ubiquitous Computing ausschließlich auf verteilten Algorithmen aufgebaut. Es gibt keine zentrale Koordination, wie es bei anderen Protokollen für ähnliche Anwendungen (z.B. Zigbee oder Bluetooth) der Fall ist. AwareCon ist vollständig implementiert und seit vielen Jahren auf der pPart Particle Computer Plattform im Einsatz.

Die Anwendungen von überlagerten Funksignalen in Sensornetzwerken sind vielfältig. Da das Thema auch in der Wissenschaft noch sehr jung ist, so ist es heute noch nicht abzusehen, ob und in welchem Anwendungsfeld sich überlagerte Funksignale durchsetzen werden. Um das breite Feld zu umreißen, das von der Anwendbarkeit von überlagerten Funksignalen aufgespannt wird, werden im Laufe der nächsten Kapitel eine Reihe von Nutzungsmöglichkeiten vorgestellt werden. Sie beziehen sich auf verschiedene Netzwerkschichten und können so auch sortiert werden. Die vorgestellten Mechanismen sind im *AwareCon*-Protokoll [12], [17] implementiert und evaluiert worden. *AwareCon* wurde dabei auf der pPart Particle Computer Plattform implementiert.

5.1 pPart Particle Computer

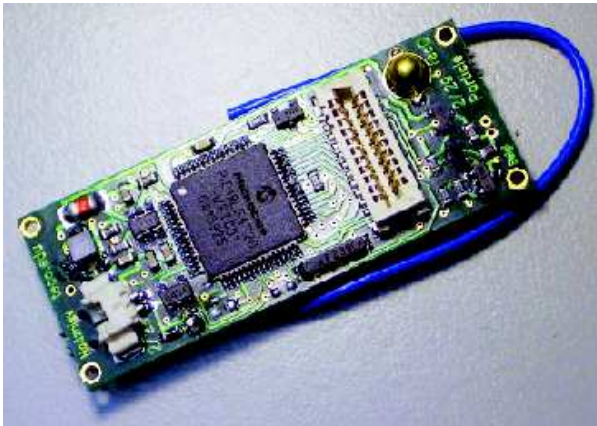


Abbildung 5.1. Ein pPart Particle Computer Sensorknoten

Die pPart Particle Computer Sensorknoten wurden am TecO in dem Projekt *Smart-Its*¹ begonnen und außerhalb dieses Projektes weiterentwickelt. Die pParts bestehen im Wesentlichen aus einem PIC18F6720 Microprozessor und einem RFM TR1001 Funktransceiver. Darüber hinaus enthalten sie Peripherie wie eine

¹<http://www.smart-its.org>

Echtzeituhr, externen Flashspeicher, Sensoren, Spannungsversorgungsschaltung, Konnektoren etc. Die pPart Particle Computer sind als Forschungsplattform entwickelt worden und sind weltweit in über 30 Forschungseinrichtungen im Einsatz. AwareCon wurde als Funkprotokoll für die pParts entwickelt und stellt die Kommunikationsbasis dar. Die technischen Eigenschaften von AwareCon sind natürlich auch stark implementierungs- und plattformabhängig. Auf den pPart Particle Computern liegt AwareCon in einer Softwareimplementierung vor, nur die Bitübertragung wird vom Transceiver übernommen. Der Transceiver bietet eine Bandbreite von 125kBit; Amplitudenmodulation ist möglich. Bei der Datenübertragung in AwareCon wird jedoch nur On-Off-Keying benutzt. Der RFM Transceiver zeichnet sich darüber hinaus durch sehr schnelle Schaltzeiten zwischen Betriebsmodi wie Empfang, Senden und passivem Stromsparmodus aus.

5.2 Das Protokoll *AwareCon*

AwareCon ist als Protokoll für Anwendungen aus dem Bereich Ubiquitous Computing entwickelt worden. Es trägt daher besondere Eigenschaften für die dort typischen Szenarien:

- Skalierbarkeit:** Mit AwareCon werden Szenarien unterstützt, die besonders hohe Gerätedichten tragen. Hierfür wurde ein spezielles Kanalzugriffsverfahren für die Anwendung auf drahtlosen Sensornetzwerken adaptiert.
- Mobilität:** Viele Anwendungsbeispiele in Ubiquitous Computing schließen Mobilität von Benutzern oder Objekten nicht aus. Dies erzeugt besondere Anforderungen an die Funkprotokolle. AwareCon unterstützt spontane Kommunikation sowie sehr kurze Synchronisationszeiten zwischen mobilen Partnern. Darüber hinaus verwendet AwareCon keine Klassifizierung von Netzwerkteilnehmern. Alle Kommunikationspartner sind prinzipiell gleichberechtigt. Dadurch unterscheidet sich AwareCon grundsätzlich von anderen bekannten Protokollen wie IEEE802.15.4 [51] oder Bluetooth.
- Energieeffizienz:** Durch die Nutzung von überlagerten Funksignalen braucht AwareCon nur 4% der Netzwerkzeit für Kanalzugriff und Synchronisationsmechanismen.

Das Protokoll AwareCon ist ein TDMA-Verfahren. Die Zeitachse wird im AwareCon-Protokoll in Zeitschlitz (Slots) unterteilt. Abbildung 5.2 zeigt den Zeitrahmen von AwareCon, der sich ständig wiederholt. Entscheidend ist für AwareCon eine sehr genaue Zeitsynchronisation zwischen den Sensorknoten.

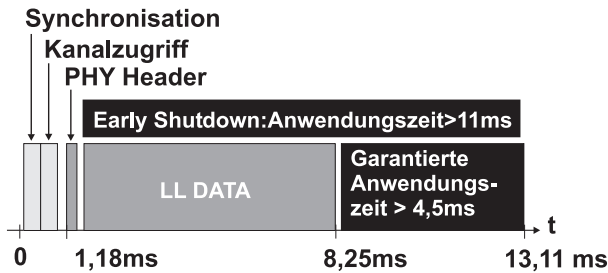


Abbildung 5.2. Der Zeitrahmen des AwareCon-Protokolls

Diese wird durch die Nutzung von überlagerten Funksignalen erreicht. Die Synchronisation findet völlig verteilt statt, es ist kein ausgezeichneter Knoten vorgesehen, der die Synchronisation vorgibt. Damit wird vor allem die Mobilität und die spontane Kommunikation unterstützt. Weiterhin kann AwareCon in der Implementierung auf der pPart Plattform mit folgenden Kernparametern charakterisiert werden:

- 4%** Overhead für Zeitsynchronisation und Kanalzugriff.
- 65%** der Prozessorzeit benötigt AwareCon bei maximaler Auslastung. Es ist somit für Single-Task-Prozessoren geeignet, die so noch mehr als 1/3 der Gesamtrechenleistung für eine zusätzliche Anwendung behalten.
- 4 μ s** Synchronisationsgenauigkeit der lokalen Uhren erreicht AwareCon zwischen *allen* Kommunikationspartnern.
- CSMA/CR.** Dieses spezielles Kanalzugriffsverfahren (*carrier sense multiple access with collision resolution*, das CSMA/CA (wie z.B. bei 802.11) um Größenordnungen schlägt und den Durchsatz auch bei mehr als 100 Geräten mit gleichzeitigem Sendewunsch hoch halten kann, wurde mit AwareCon zu ersten Mal in der drahtlosen Welt eingeführt. In Abbildung 5.3 ist der Kanalzugriff zwischen CSMA/CA und AwareCon verglichen worden. Man sieht, dass AwareCon deutlich geringere Kollisionsraten erzeugt.

Der gesamte Ablauf des Protokolls ist in Anhang B zu finden. In den nun folgenden Kapiteln werden zwei der zentralen Elemente von AwareCon vorgestellt: *Zeitsynchronisation* und *Kanalzugriff*. Für beide Mechanismen werden überlagerte Funksignale eingesetzt und so herausragende Ergebnisse erzielt.

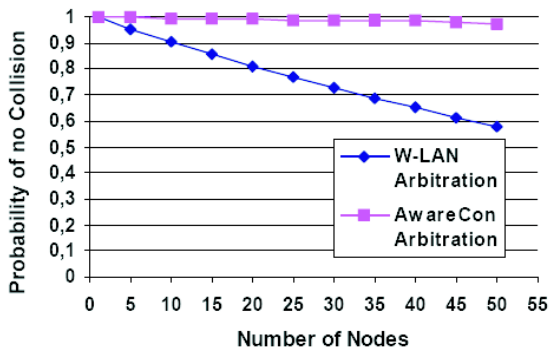


Abbildung 5.3. Kanalzugriff im AwareCon-Protokoll mit CSMA/CR. Aus [17, Abbildung 6]

5.3 Zeitsynchronisation in AwareCon

Überlagerte Funksignale für die Zeitsynchronisation zu verwenden, ist in der Literatur bereits mehrfach diskutiert worden. Die Arbeiten reichen von Ultrabreitband ([28]) über FSK-Systeme ([34]) bis zu gekoppelten Oszillatoren ([27]). Das Zeitsynchronisationsverfahren in AwareCon basiert auf dem verbreiteten Ansatz von periodischen „Beacons“. Diese Beacons sind Datenpakete, die zwischen Knoten ausgetauscht werden und einen Zeitreferenzpunkt enthalten. So kann jeder Knoten, der ein solches Paket empfängt, seine lokale Uhr nachstellen und sich so *resynchronisieren*. Der Zeitsynchronisationsmechanismus in AwareCon fußt auf einer Idee, die von M. Ringwald als BitMAC [67] veröffentlicht wurde. BitMAC definiert ein Protokoll, das auf der Übertragung einzelner Bits auf der Zugriffsschicht beruht. In BitMAC werden Präambeln und weitere typische Paketteile bewusst weggelassen, um eine effiziente Übertragung von einzelnen Bits zu erreichen. Für die Zeitsynchronisation wird nun eine gewisse Sequenz von Bits versendet. Dies kann (nach erfolgreicher Synchronisation) auch simultan von mehreren Quellen ausgehen: Die Überlagerung auf dem Kanal soll in BitMAC problemlos möglich sein. BitMAC verwendet als Arbeitshypothese hierfür die Annahme, dass der Funkkanal wie eine „Oder“-Verknüpfung arbeitet. Ein Beispiel für BitMAC ist nun in Abbildung 5.4 zu sehen. Zwei Quellen senden simultan

die gleiche Sequenz aus. Durch die leichte Zeitverschiebung und die Annahme der „Oder“-Charakteristik des Kanals ergibt sich ein erfolgreicher Empfang der Sequenz. Bitweise Übertragungen wurden auch schon vorher im Protokoll SDJS [45] verwendet. Dieses Protokoll wird in Kapitel 7 genau erläutert werden. Weiterhin verwenden BitMAC stets eine zentrale Instanz, von der die Zeitsynchronisation ausgeht.

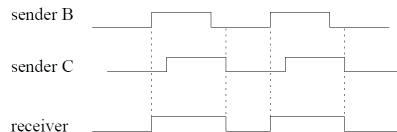


Abbildung 5.4. Identische Aussendungen von zwei Sendern mit geringen Zeitsynchronisationsungenauigkeiten. Der Empfänger wird leicht gestreckte 1-Bits und leicht verkürzte 0-Bits empfangen (aus [67, Abbildung 5, Seite 6]).

In AwareCon wird die Idee zur Zeitsynchronisation durch simultane Emission von einer bestimmten Bitsequenz aufgegriffen und unter dem Aspekt der überlagerten Funksignale genauer betrachtet. Die Erweiterungen von BitMAC für die Zeitsynchronisation sind:

Praktische Betrachtung der Überlagerung: Die Annahme eines „Oder“-Kanals auf der Zugriffsebene wird durch das Modell der überlagerten Funksignale ersetzt. Die Signale werden nach ESK in Kapitel 3.5 geformt.

Verteilter Fall: AwareCon soll keine zentralistische Struktur beinhalten.

Die Zeitsynchronisation mit überlagerten Funksignalen wurde im Protokoll AwareCon ([12], [17]) implementiert und stellt dort die Basis für das TDMA Protokoll. Die nächsten Abschnitte befassen sich nun mit folgenden Einzelaspekten:

Aufbau der Synchronisation: Der Ablauf, wie Synchronisationen aufgebaut und erhalten werden. Hier ist besonders der vollständig verteilte Algorithmus wichtig, da AwareCon keine Netzwerkknoten mit besonderer Funktionalität vorsieht.

Signalisierung: Wie wird eine bestehende Synchronisation zwischen Knoten erhalten? Auch hier ist es wichtig, dass die verteilte Synchronisation in gleichem Maß von allen Teilnehmern unterstützt wird. Bei der Signalisierung kommen überlagerte Funksignale zum Einsatz.

Evaluation und Genauigkeit: Welche quantitativen Genauigkeiten werden mit dem AwareCon-Protokoll erreicht?

Verteilte Synchronisation: Der verteilte Fall ohne zentrale Synchronisationsquelle beinhaltet einige theoretische Hürden, für die hier Lösungen diskutiert werden.

5.3.1 Aufbau der Synchronisation

Die Synchronisation im AwareCon-Protokoll ist speziell für Szenarien aus dem Bereich Ubiquitous Computing entworfen worden. Deshalb ist es wichtig, dass das Synchronisationsverfahren verteilt und ohne Verwendung eines ausgezeichneten Knotens geschieht. Jeder Knoten ist in gleicher Weise an der Synchronisation beteiligt und auch verantwortlich. Sobald zwei oder mehr Knoten in gegenseitiger Reichweite sind, können sie sich zueinander synchronisieren. Dazu verwenden sie den Ablauf, der in Abbildung 5.5 dargestellt ist. Hier wird vor allem zwischen den Zuständen *synchronisiert* und *nicht synchronisiert* unterschieden. Knoten, die synchronisiert sind, nehmen in jedem Zeitrahmen an der Synchronisation teil; sie senden entweder ein Synchronisationssymbol aus, oder versuchen, ein von anderen Knoten gesendetes zu empfangen. Somit nimmt jeder Knoten aktiv und zufällig gesteuert an der Synchronisation teil und hält durch dieses verteilte Verfahren die Synchronität zwischen den Funkknoten aufrecht. Die Wahrscheinlichkeiten des Sendens und Empfangens sind entscheidende Parameter im System. Wenn in einem Zeitslot z.B. alle Knoten ein Synchronisationssymbol senden, so entsteht dadurch keine Resynchronisation. Nur Knoten, die Synchronisationssymbole *empfangen*, können ihre lokale Uhr entsprechend ihrer Netzwerkpartner nachstellen. Um hier ein sinnvolles Gleichgewicht zwischen Senden und Empfangen zu erhalten, wird ein zusätzlicher Schätzmechanismus eingesetzt, der die Anzahl der anwesenden Netzwerkpartner schätzt. Wie diese Schätzung funktioniert, wird in Kapitel 7 detaillierter erklärt. Basierend auf der Anzahl der Geräte entscheiden sich dann die Funkknoten zufällig, ob sie nun ein Synchronisationssymbol senden oder versuchen eines zu empfangen. In Szenarien, wo sehr wenige Funkknoten anwesend sind, müssen die Knoten eine höhere Senderate der Synchronisationssymbole haben, damit keine zu langen Zeiträume entstehen, in denen kein Knoten ein Synchronisationssymbol sendet.

Die Sendewahrscheinlichkeit für Synchronisationssymbole ist ein Parameter, der auch verwendet werden kann, um Extremfälle zu parametrisieren. Setzt man z.B. die Senderate eines Knotens fest auf 100% und die Senderate aller anderen

auf 0%, so wird aus der verteilten Synchronisation eine zentralisierte. Der 100%-Knoten ist die Synchronisationsvorgabe für alle anderen. Dies entspricht einer Master-Slave-Synchronisation, wie sie in vielen bekannten Protokollen eingesetzt wird.

Wenn nun mehrere Knoten synchronisiert sind, und sich gleichzeitig dazu entscheiden, ein Synchronisationssymbol auszusenden, so muss ein Verfahren gefunden werden, dass den Kanalzugriff für solch einen Fall klärt. Im Falle von AwareCon wird der Kanalzugriff nicht geklärt, es wird stattdessen vorgeschlagen, dass gleichzeitig sendeinteressierte Knoten auch tatsächlich gleichzeitig Synchronisationssymbole aussenden dürfen. Diese überlagern sich dann im Kanal. Durch die Nutzung von überlagerten Funksignalen bleibt die Summe der ausgesandten Signale jedoch noch als Synchronisationssymbol erkennbar.

5.3.2 Signalisierung

In Abbildung 5.6 ist der Synchronisationsmechanismus zwischen drei Knoten dargestellt. Knoten A und B sind bereits synchronisiert und Knoten C noch nicht. Teilnehmer, die bereits synchronisiert sind, senden Synchronisationssymbole aus und Teilnehmer, die noch nicht synchronisiert sind, versuchen diese Synchronisationssymbole zu empfangen. Beim Versenden von Synchronisationssymbolen werden Überlagerungen genutzt, um eine automatische räumliche Mittelung zu erzeugen und die Sendeleistung zu erhöhen. In Abbildung 5.6 senden Knoten A und B gleichzeitig eine Synchronisation aus. Sie überlagern also ihre Signale auf dem Kanal. Diese Überlagerung wird von Knoten C gesehen. Da die Synchronisation — wie in Abbildung 5.6 mit den gestrichelten Linien angedeutet — nicht perfekt ist, kommt es im Empfänger (Knoten C) zu einer „Verschmierung“ der Signalgrenzen. Deshalb kann Knoten C nicht die Signalgrenzen verwenden, um den Synchronisationszeitpunkt zu bestimmen. Stattdessen wird während der Präambel des Synchronisationssymbols die Mitte der Signale ermittelt. Diese Mitte enthält auch automatisch die räumliche Mittelung der ausgesandten Signale der verschiedenen Sendestationen. Die minimalen Restverschiebungen zwischen Teilnehmern werden so automatisch auf dem Kanal gemittelt. Zum Zeitpunkt „Zeitreferenz“ werden dann alle Uhren auf einen gemeinsamen Wert nachgestellt.

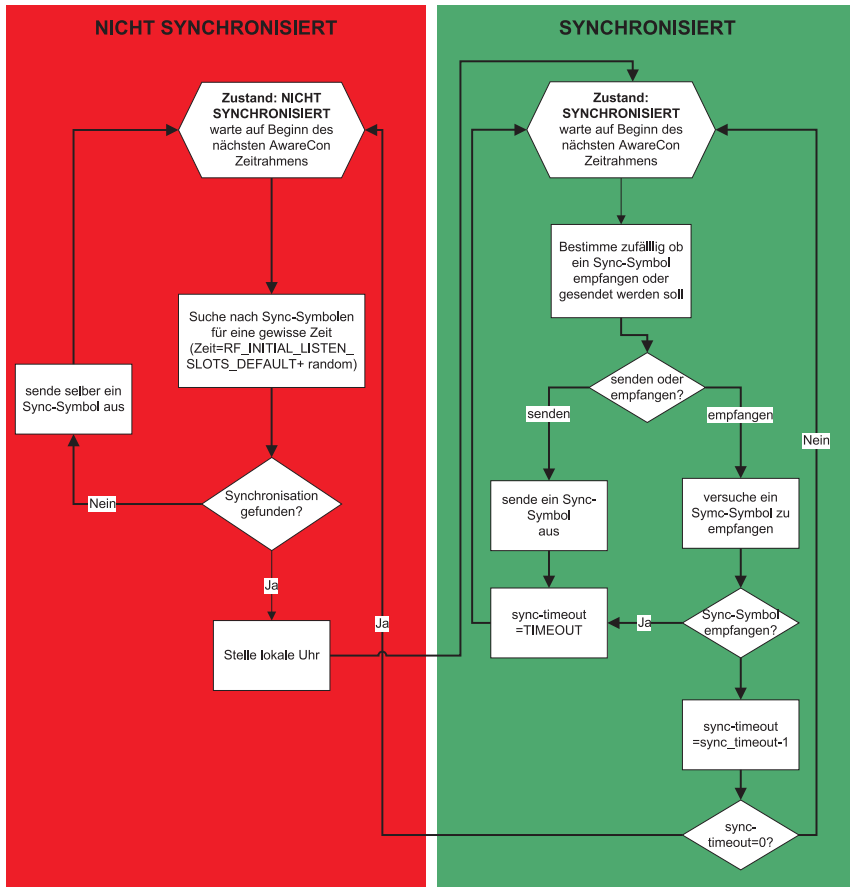


Abbildung 5.5. Der Ablauf der Synchronisation im AwareCon-Protokoll

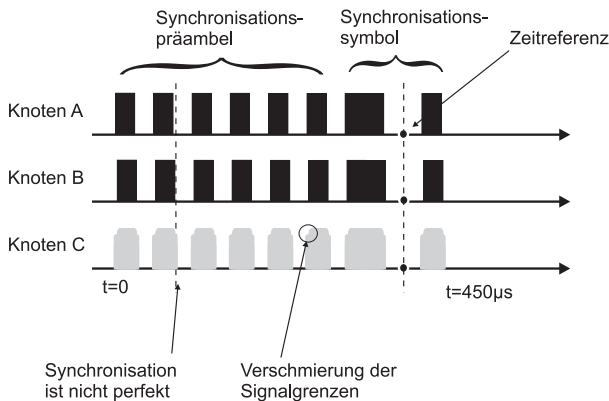


Abbildung 5.6. Der Synchronisationsmechanismus des AwareCon-Protokolls mit überlagerten Funksignalen

5.3.3 Evaluation der Synchronisation in AwareCon

Mit AwareCon auf der pPart Plattform wurden einige Testszenarien aufgebaut, um die Performance des vorgeschlagenen Mechanismus zu beurteilen. Da AwareCon speziell auch für mobile Szenarien entworfen wurde, lag besonderes Augenmerk hier auf der Geschwindigkeit, mit der ein Knoten sich mit anderen synchronisieren kann. In Abbildung 5.7 sind die Synchronisationszeiten für zwei experimentelle Beispiele dargestellt. Der Versuchsaufbau wurde in einer Büroumgebung durchgeführt. Dazu wurden pPart Particle Computer an eine schaltbare Versorgungsspannung angeschlossen. Nach dem Einschalten der Versorgung startet das Protokoll AwareCon und versucht, sich mit anderen AwareCon-Funkknoten zu synchronisieren. Die zwei Versuche unterscheiden sich in der Anzahl der teilnehmenden Funkknoten. Beim ersten Versuch wurden zwei pParts Particle Computer gleichzeitig gestartet. Beim zweiten Versuch besteht bereits eine Synchronisation von zwei pParts und es wird ein weiteres pPart aufgeschaltet. Gemessen wurde jeweils die Zeitdauer vom Einschalten des pParts, bis der Status *synchronisiert* aus Abbildung 5.5 erreicht war.

Es ist klar zu sehen, dass die Synchronisation zu einem schon bestehenden Netzwerk deutlich schneller durchgeführt werden kann als die zu nur einem Partner. Dies liegt daran, dass bei einer bestehenden Synchronisation der neue Kno-

ten lediglich die Synchronisationssymbole der anderen Partner finden muss. Es existieren nur selten Pausen ohne Synchronisationssymbol, die länger als ein AwareCon-Zeitschlitz (13ms) sind. Die mittlere Dauer zum Synchronisieren auf ein bestehendes Netz wurde in diesem Szenario experimentell mit $12.6ms$ bestimmt. Wenn nur zwei Funkknoten eine gemeinsame Synchronisation etablieren wollen, so ist der Prozess etwas komplizierter. Wie in der Flussgraphik 5.5 zu erkennen, existieren längere Wartezeiten, in denen Funkknoten nur auf Empfang geschaltet sind und kein Synchronisationssymbol ausgesendet wird. In den Experimenten betrug hier die mittlere Synchronisationszeit zwischen zwei Partnern $40ms$. Die Streuung der benötigten Synchronisationszeit ist, wie in Abbildung 5.7 zu sehen, jedoch recht breit. Das liegt daran, dass die Synchronisation auf Zufallsprozessen basiert, und es deshalb möglich ist, dass beide Partner z.B. sehr lange nur den Kanal beobachten, was zu einer langen Synchronisationsverzögerung führt, da von keinem der Partner ein Symbol gefunden werden kann. Es ist jedoch auch möglich, dass sich einer der Partner sehr schnell zum Aussenden eines Synchronisationssymbols entscheidet und so die Synchronisation vorgibt, die der andere Partner dann bestätigt und so schnell eine gemeinsame Synchronisation etabliert wird.

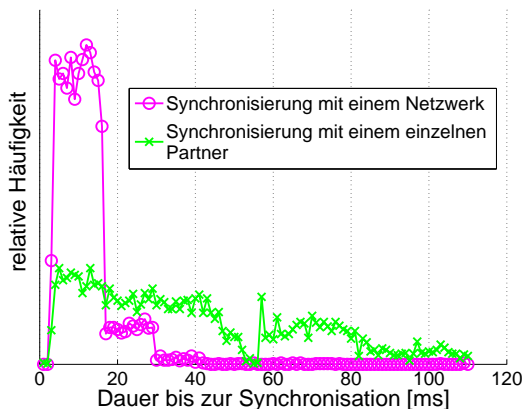


Abbildung 5.7. Die Verzögerungszeit, bis Teilnehmer im AwareCon-Protokoll synchronisiert sind

5.3.4 Verteilte Synchronisation

Für das Synchronisationsprotokoll für AwareCon müssen für den verteilten Fall noch zwei zentrale Probleme bearbeitet werden. Diese sollen nun als *Synchronisationsketten* und *konkurrierende Inseln* bezeichnet werden. Diese Probleme sind in der Netzwerktheorie für Sensornetzwerke bekannt und bestehen prinzipiell für alle verteilten Zeitsynchronisationsmechanismen wie z.B. T-MAC [16] oder S-MAC [85]. Hierfür existieren unterschiedliche Lösungsansätze für verschiedene Systemvoraussetzungen.

Synchronisationsketten

Das Problem der Synchronisationsketten entsteht dann, wenn eine Zeitreferenz durch eine Multihop-Kommunikation einen Kreisschluss erzeugt. Dabei kann sich die Zeitsynchronisation durch Fehlerfortpflanzung so weit verschieben, dass sie beim Endpunkt (der in diesem Beispiel zugleich der Startpunkt der Synchronisation war) außerhalb des Toleranzbereiches liegt. Die Bits für das Synchronisationssymbol sind dann so weit gegeneinander verschoben, dass es nicht mehr zu sinnvoller Überlagerung kommen kann. Diesem Effekt, der nur auf der Genauigkeit der Zeitsynchronisation beruht, kann auf verschiedene Weise begegnet werden: Es können z.B. die Bitlängen auf die maximale Verschiebung ausgelegt werden. Je größer diese maximale Verschiebung sein kann, desto länger müssen die Bits in der Implementierung sein, um Laufzeiten und summierte Fehler zu tolerieren. Hierfür muss allerdings eine Annahme über die maximale Verschiebung, also auch die maximale Anzahl von Hops in einem Netzwerk getroffen werden.

Konkurrierende Inseln

Der Fall von konkurrierenden Inseln tritt dann auf, wenn sich zwei mobile, synchronisierte Netze begegnen. Die Knoten an den Berührungspunkten der Netze können nicht mehr eindeutig entscheiden, welchem Netz sie angehören und welchem Zeitschema sie folgen sollen. In S-MAC wird hier vorgeschlagen, dass Grenzknoten einfach beide Zyklen annehmen. Ein anderer Vorschlag wäre, mehrere Funkkanäle zu verwenden, um so den Konflikt von zwei Zeitschemen im gleichen Band zu vermeiden. Prinzipiell sind für diese Problematik drei Aspekte relevant:

Konflikterkennung: Knoten müssen feststellen, dass zwei konkurrierende Zeitschemen existieren. Hierfür müssen sie spezielle Entdeckungsmechanismen verwenden.

Konfliktkommunikation: Ein entdeckter Konflikt muss an Netzwerkpartner kommuniziert werden. Hier bieten sich Signalisierungen im Band, wie z.B. *Jamming*, oder außerhalb des Bandes, wie z.B. *busy-tones*, an.

Konfliktauflösung: Ein bestehender Konflikt muss aufgelöst werden. Dies geschieht z.B. über ein Präferenz- und Prioritätsmaß auf den Knoten. Eine verbreitete Lösung ist die Verwendung der Identifikation als eindeutiges Maß.

Die beiden letzten Punkte sind besonders dann relevant, wenn — im Gegensatz zu S-MAC — nach einer globalen Lösung gesucht werden soll. Eine globale Lösung würde bedeuten, dass zwei synchronisierte Netze, die sich begegnen, nicht nur in den Berührungspunkten, sondern für alle Knoten eine neue, gemeinsame Synchronisation anstreben. Eine Möglichkeit wäre hier (wie in [55]) über eine globale Präferenz (z.B. die Identifikationsnummer der Knoten) eine neue Synchronisation auszulösen. Nimmt man, wie für die Synchronisationsketten, auch z.B. eine maximal mögliche Anzahl von Hops in einem Netz an, so könnte die Identifikationsnummer eines Knotens mit Hilfe des Synchronisationssymbols propagiert werden. Die größte Nummer würde in der Resynchronisation Vorrang haben und würde von Partnerknoten weiterpropagiert. Damit Knoten im verteilten Fall nicht die Identifikationsnummer anderer Knoten in einem „Ping-Pong“-Prozess dauerhaft weiterpropagieren, könnte das bekannte Verfahren einen zusätzlichen Wiederholungszähler (time to live TTL) einsetzen. Nach einer gewissen Zahl von Hops würde diese höchste Identifikationsnummer dann nicht weiter propagiert. Mit der vorher angenommenen maximalen Hopzahl im Netz ergibt sich so eine praktikable Lösung.

5.4 Der Kanalzugriff in AwareCon

Das Kanalzugriffsverfahren, das in AwareCon die Vergabe der Zeitschlitze regelt, ist aus der Literatur als CSMA/CR bekannt. Es bildet die Basis des Kanalzugriffs für einen der meisteingesetzten Industriebusse: dem CAN (controller area network) Bus [37]. Es wird dort typischerweise eingesetzt, um eine feine Auflösung von Priorisierungen von Nachrichten zu ermöglichen. Bei bekannten und a-priori vergebenen Prioritäten, die für alle denkbaren Nachrichten unterschiedlich sind,

kann CAN eine 100%ige Kollisionsfreiheit garantieren. Im Zugriffsverfahren basiert es auf der Verwendung von dominanten und rezessiven Bits und wird „non-destructive bit-wise arbitration“ genannt. In Abbildung 5.8 sieht man den konkur-

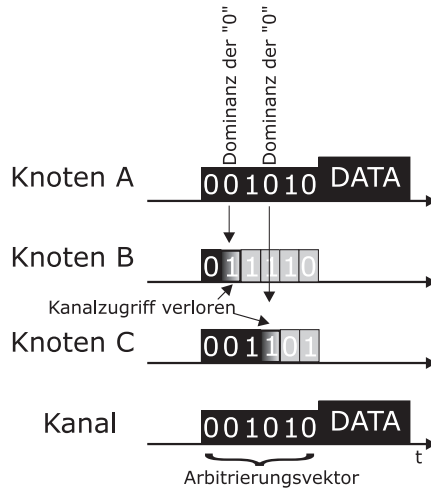


Abbildung 5.8. Die Darstellung der „non-destructive bit-wise arbitration“

rierenden Kanalzugriff zwischen den Knoten A, B und C. Für den Kanalzugriff wählen die beteiligten Knoten eine binäre Zufallszahl, den *Arbitrierungsvektor*. Anschließend werden diese Zufallszahlen bitweise auf den Kanal übertragen, und an jeder Bitstelle eine Entscheidung getroffen, welcher Knoten weiterhin versuchen darf, auf den Kanal zuzugreifen. Die Bits der Arbitrierungsvektoren werden als dominante und rezessive Signale auf der physikalischen Schicht realisiert, so dass die „0“ ein dominantes und die „1“ ein rezessives Signal erhält. Für kabelgebundene Systeme wie z.B. CAN wird dies durch einen sogenannten tri-state-Bus auf einer Zweidrahtleitung realisiert. Der Bus ist im Ruhezustand über hochohmige Widerstände auf eine bestimmte Spannung gezogen. Stationen, die ein dominantes Signal verschicken, schließen den Bus kurz und erzeugen so ein 0V Signal auf dem Bus. Wenn dies mehrere Stationen tun, so bleibt der Bus bei 0V. Die Stationen, die ein rezessives Bit zu verschicken haben, schließen den Bus nicht kurz, sondern messen, ob der Bus auf Spannung oder 0V ist. Ist er auf 0V, so haben sie die Arbitrierung zu diesem Zeitpunkt verloren, weil sie selbst ein rezessives

Bit hatten und eine andere Station ein dominantes hatte. Auf diese Weise wird Bit für Bit entschieden, welche Stationen noch an dem Kanalzugriff teilnehmen. In Abbildung 5.8 gewinnt Knoten A den Kanalzugriff, weil er an den Bitpositionen 2 und 4 ein dominantes Bit hat, wo Knoten B und C jeweils rezessive Bits haben.

Für den Einsatz dieses Zugriffsverfahrens für drahtlose Sensornetzwerke muss man zunächst einen technischen Unterschied zwischen kabelgebundenen und kabellosen Systemen feststellen: Die Verwendung von dominanten und rezessiven Bits durch einen tri-state-Bus mit hochohmigen Widerständen ist so nicht direkt darstellbar. Um die gleiche Funktionalität für ein Funksystem zu erreichen, sollen hier überlagerte Funksignale eingesetzt werden:

5.4.1 Überlagerte Funksignale als virtueller tri-state-Bus

Um dominante und rezessive Bits auf dem Funkkanal zu realisieren, wird nun vorgeschlagen, eine spezielle Signalisierung aus überlagerten Funksignalen einzusetzen, die die Funktionalität eines kabelgebundenen tri-state-Busses erfüllen. Folgende Analogien werden hierfür erkannt:

1. Das Versenden eines dominanten Bits entspricht bei der kabelgebundenen Implementierung einem Kurzschluss des Übertragungskabels. Für Funksysteme wird hier das Versenden von *Jamsignalen* vorgeschlagen.
2. Das Versenden eines rezessiven Bits bedeutet bei der kabelgebundenen Implementierung, dass der entsprechende Transceiver kein Signal auf den Bus legt, also einen hochohmigen Zustand annimmt. Stattdessen detektiert er den Kanal und stellt fest, ob es ein Signal auf dem Bus gibt. Für Funksysteme soll hier einfach der Transceiver auf Empfang geschaltet werden und eine Signaldetektion durchgeführt werden.

Da die Anwendung von *Jamsignalen* in dieser Arbeit ein wichtiges technisches Hilfsmittel ist, werden sie für den weiteren Verlauf nun definiert:

Definition 5.1 (Jamsignale) *Jamsignale sind zufällige, bandbegrenzte Zeitsignale. Jamsignale werden nur durch ihre Leistung, die Bandbreite und die Signaldauer festgelegt. Jamsignale können unterschiedlich generiert werden. Eine Möglichkeit für einfache Funktransceiver ist es, den Träger für die gegebene Signaldauer eines Jamsignals einzuschalten. Eine andere Möglichkeit ist es, ein bandbegrenztes Rauschsignal auf den Träger zu modulieren und auszusenden. Jamsignale*

werden im Empfänger nur durch ihre Dauer und Leistung charakterisiert. In einfachsten Fall werden im Empfänger nur zwei Leistungsbereiche unterschieden: „Rauschleistung“ und „mehr als Rauschleistung“. Eine solche Detektion würde dann einem carrier sense-Mechanismus auf der physikalischen Schicht entsprechen.

Netzwerkteilnehmer, die einen virtuellen tri-state-Bus auf der Funknetzwerkebene nutzen wollen, müssen einige Anforderungen erfüllen. Diese sind:

1. Die Möglichkeit, Jamsignale zu verschicken
2. Die Möglichkeit, Jamsignale von anderen zu erkennen
3. Eine genaue Zeitsynchronisation zwischen Teilnehmern
4. Methoden für zuverlässigen Rundruf
5. Methoden für Kanalreservierung

Diese Anforderungsliste erscheint zunächst umfangreich, doch sind diese Voraussetzungen typische Eigenschaften von Funkeinheiten und erfordern normalerweise keinerlei zusätzliche Hardware-Anpassung. Unter einem Jamsignal wird hier z.B. einfach die Trägerfrequenz verstanden. Versenden eines Jamsignals würde also bedeuten, das Trägersignal einzuschalten. Zusätzlich sollte man Signaturen zur Bandspreizung verwenden, wie sie in Kapitel 3.1.3 vorgeschlagen wurden, um die Erkennung der Signale in den Empfängern zu verbessern. Für den Empfang wird die Erkennung von fremden Trägern typischerweise als CCA (clear channel assessment)-Mechanismus bereits von der Hardware der Transceiver bereitgestellt. Durch den in Abschnitt 5.3.1 erklärten Ablauf des Verfahrens ist auch die Zeitsynchronisation kein Problem. Für andere, einfache Funksignale kann sie auch mit Hilfe des Startpakets etabliert werden. Somit erfordert dieses Verfahren keine zusätzlichen Mechanismen, die Standardhardware und Protokollimplementierungen nicht sowieso schon mit sich bringen.

Der Überlagerungsprozess basiert im Wesentlichen auf zwei Mechanismen, nämlich dem *Versenden von Jamsignalen* und dem Entdecken von *Jamsignalen*. Letzteres wird z.B. über den carrier sense-Mechanismus der Funkimplementierung realisiert. Hier werden typischerweise Entdeckungswahrscheinlichkeiten von über 90% erwartet [35, Kap. 14.6.15.3]. Typische Fehler bei der Entdeckung wie Fehlalarme und schlechte Erkennung werden im Folgenden nicht in die Modelle aufgenommen. Sie sind eine Frage der Implementierung (insbesondere der Hardware und der Detektion). Darüber hinaus müssen die Teilnehmer sehr schnell zwischen Sende- und Empfangszustand wechseln können. Auch die Erkennung von Signalen muss innerhalb der Slotzeit gelingen. Hierdurch entsteht

eine untere Grenze für die Wahl der Slotzeit, die im Wesentlichen von den Leistungsmerkmalen der physikalischen Schicht abhängt. Wenn man die Parameter der physikalischen Schicht von IEEE 802.11a [35] nimmt, so erfährt man für die Slotzeit eine untere Grenze:

$$\begin{aligned} t_{Slot} &= \max(t_{rx/tx}, t_{tx/rx}) + 2 \cdot t_{channel\,delay} + t_{carrier\,sense} \\ &= \max(2\mu s, 12\mu s) + 2 \cdot 1\mu s + 4\mu s \\ &= 18\mu s \end{aligned} \tag{5.1}$$

Mit einem Funkfrontend, das IEEE 802.11a unterstützt, kann also der Überlagerungsprozess mit $18\mu s$ Slotzeit ($=55\text{kBit/s}$) durchgeführt werden. Diese theoretische, minimale Slotzeit muss eventuell noch weiter entspannt werden, um Synchronisationsschwächen und Quarzdrift zu kompensieren. Auf der pPart Plattform ist diese Slotzeit z.B. $32\mu s$. Mit den vorgestellten Analogien in einem Funksystem mit überlagerten Funksignalen kann also das Zugriffsverfahren CSMA/CR für drahtlose Systeme eingesetzt werden.

5.4.2 Arbitrierung in AwareCon

AwareCon nutzt den CSMA/CR-Mechanismus, um den Kanalzugriff zu organisieren. In jedem 13ms -Zeitrahmen von AwareCon kann nur jeweils ein Funkknoten ein Paket verschicken. Der Zugriff wird dabei zufällig auf alle Teilnehmer mit Sendewunsch verteilt. Hierfür wählen die Teilnehmer einen zufälligen Zugriffsvektor einer Länge von 10Bit. Dieser Vektor ist der Arbitrierungsvektor. Mit dem CSMA/CR-Verfahren wird dann der Kanalzugriff geklärt. Auf diese Weise kann AwareCon sehr große Teilnehmerzahlen handhaben, ohne übermäßige Kollisionsraten zu erzeugen. Obwohl der Kanalzugriff in AwareCon standardmäßig einen fairen Zugriff für alle Teilnehmer vorsieht, ist CSMA/CR hervorragend geeignet, um Priorisierungen auf Kanalzugriffsebene zu realisieren. Dies wird nun im folgenden Kapitel erläutert.

5.4.3 TOMAC

Mit TOMAC [43] wurde das Zugriffsverfahren CSMA/CR für eine Anwendung in drahtlosen Sensornetzwerken angepasst. TOMAC kann Nachrichtensortierung auf der Kanalzugriffsschicht realisieren und so z.B. Echtzeitsysteme unterstützen.

Das zeitliche Sortieren von Nachrichten ist ein bereits breit diskutiertes Thema. Dennoch verwenden bekannte Verfahren keine Mechanismen auf der Zugriffsschicht, um Nachrichtensortierung durchzuführen. In diesem Zusammenhang soll nun das CSMA/CR-Verfahren mit der Erweiterung für drahtlose Systeme verwendet werden, um eine Nachrichtensortierung zu erreichen. Die Sortierung wird über Priorisierungen von Kanalzugriffen geregelt. Die Priorisierungen von Nachrichten werden während der Laufzeit dynamisch bestimmt. Der wesentliche Eingabeparameter ist hier die Wartezeit. So kann mit Hilfe von CSMA/CR garantiert werden, dass länger wartende Pakete eine höhere Priorität haben als kürzer wartende und deshalb im konkurrierenden Kanalzugriff bevorzugt werden. Die Wartezeit kann hierbei z.B. ab dem Feststellen eines physikalischen Ereignisses gemessen werden. So lässt sich sicherstellen, dass physikalische Ereignisse, die als erste entdeckt wurden, auch als erste auf dem Kanal übermittelt werden.

Alle Nachrichten werden mit einer eigenen Priorisierung r_i versehen. Mit zunehmendem Zeitverlauf wird diese Priorität automatisch mit der Steigerung dr_i/dt erhöht. Der Wert der Priorität beginnt bei einem minimalen Wert r_{min} und ist begrenzt durch r_{max} . In Abbildung 5.9 ist der Ablauf der Kommunikation mit TOMAC beispielhaft dargestellt. Zur Zeit t_1 hat Knoten A einen Sendewunsch. Das Medium ist frei, deshalb kann A seine Nachricht sofort verschicken. Zu den Zeiten t_2 und t_3 erzeugen Knoten B und C ebenfalls einen Sendewunsch. Da das Medium nicht frei ist, muss die Sendung verschoben werden. Das Prioritätssystem von TOMAC kommt zum Tragen. Während die beiden Knoten darauf warten, dass das Medium wieder frei wird, erhöhen sich die Prioritäten ihrer wartenden Pakete automatisch und in gleicher Weise (dr/dt ist ein systemweit konstanter Wert). Zur Zeit t_4 , wenn der Kanal wieder frei ist, wird der Zugriff auf den Kanal zwischen Knoten B und C ausgehandelt. Dafür wird das vorgestellte CSMA/CR-Verfahren eingesetzt. Anstelle einer Zufallszahl für den Kanalzugriff wird jetzt die Priorität einer Nachricht verwendet, um den Kanalzugriff zu klären. Die Priorität von B ist höher und somit gewinnt B die Arbitrierung und kann seine Nachricht zuerst verschicken. Auf diese Weise wird eine Sortierung der Nachrichten entsprechend ihrer Wartezeit realisiert. Es stellt sich nun die Frage, warum man für ein solches System nicht einfach ein CSMA-Verfahren einsetzt und die Priorität über Wartezeiten bei Kanalzugriff organisiert. So würde ein schon länger wartender Knoten, wenn der Kanal frei wird, früher als ein anderer versuchen, auf den Kanal zuzugreifen und so auch eine Priorisierung durchsetzen können. CSMA/CR hat jedoch durch seine bit-weise Arbitrierung einen wesentlich größeren Raum an möglichen Priorisierungen. Legt man z.B. 100 slots für

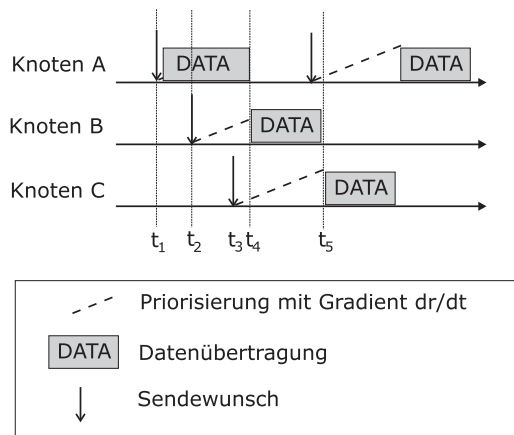


Abbildung 5.9. Die Priorität einer wartenden Sendung nimmt mit zunehmender Wartezeit zu.

ein CSMA-System fest, so können nur 100 verschiedene Anfangszeiten für Sender unterschieden werden, und somit auch nur 100 verschiedenen Prioritäten. Mit CSMA/CR hingegen sind es 2^{100} . So kann also die Zeitaufösung wesentlich feingranularer dargestellt werden und größere Zeitspannen können abdeckt werden.

5.5 Zusammenfassung

Mit AwareCon finden überlagerte Funksignale zum ersten Mal Anwendung in einem Funkprotokoll. AwareCon nutzt überlagerte Funksignale gleich mehrfach: in der Synchronisation und für den Kanalzugriff. Die Vorteile, die AwareCon dadurch gegenüber anderen Protokollen erreicht, sind signifikant. So kann z.B. eine Zeitsynchronisation von $4\mu s$ zwischen allen Knoten erreicht werden, die auch experimentell nachgewiesen wurde. Die Zeitsynchronisation im verteilten Fall kann jedoch durch weitere Mechanismen noch verbessert werden, um die Probleme der Inseln und Synchronisationsketten effizient aufzulösen. Basis des Zugriffsmechanismus ist der virtuelle tri-state-Bus, der durch überlagerte Funksignale ermöglicht wird. Mit überlagerten Funksignalen ist das Versenden von dominanten und rezessiven Bits auch in Funkprotokollen möglich, und so kann AwareCon eine

bitweise Arbitrierung für den Kanalzugriff verwenden. Diese zeichnet sich durch exzellente Skalierbarkeit und niedrige Kollisionsraten aus. Die Erweiterung der Arbitrierung im Zugriffsprotokoll TOMAC erzeugt sogar ein Echtzeitsystem, das Nachrichten auf der physikalischen Schicht nach ihrer Wartezeit sortieren kann. Der jahrelange Einsatz von AwareCon auf der pPart Particle Computer Plattform zeigt die Stabilität und praktische Nutzbarkeit von überlagerten Funksignalen in einem Protokoll für Sensornetzwerke.

6 Erhöhung der Erreichbarkeit und Zuverlässigkeit mit überlagerten Signalen

Drahtlose Sensornetzwerke werden in sehr unterschiedlichen Umgebungen eingesetzt. In Installationen mit einer geringen Anzahl von Sensorknoten pro Fläche werden traditionellerweise Multihop-Mechanismen eingesetzt, um Informationen von einer Quelle zu einem Ziel zu transportieren. Hierbei werden Informationen in kleine Einheiten zerlegt und von Funkknoten zu Funkknoten weitergereicht. Ein Problem bei dieser Vorgehensweise ist, dass lokale Unterbrechungen von Funkverbindungen dazu führen können, dass das gesamte Netz in disjunkte Teile partitioniert wird, weil entscheidende Kommunikationsbrücken unterbrochen sind. Mit der Anwendung von überlagerten Funksignalen ist man jedoch in der Lage, diese entscheidenden Kommunikationsbrücken wieder herzustellen und die vorher separierten Teile des Netzes wieder zusammenzuführen. In dem folgenden Kapitel wird gezeigt, dass man mit dem Einsatz überlagerter Funksignale die Konnektivität zwischen Funkknoten eines Netzes bis zu 50% erhöhen kann. Gleichzeitig ist es auch möglich, mit einer bis zu 30% geringeren Anzahl von Knoten die gleiche Fläche auf Sensorereignisse zu überwachen. Um überlagerte Funksignale in solchen Szenarien einzusetzen, müssen Koordinationsmodelle erarbeitet werden. Im folgenden Kapitel werden nun drei verschiedene Koordinationsmodelle vorgestellt und verglichen.

Drahtlose Sensornetzwerke gewinnen mehr und mehr Bedeutung als ein Instrument für feingranulare Beobachtung einer Umgebung oder physikalischer Ereignisse. Um ein Sensornetz zu bilden, werden Sensorknoten in einem Bereich ausgebracht, um dort die interessanten Messgrößen aufzunehmen. Solche Netze können aus bis zu mehreren hundert Knoten bestehen. Die Installation eines solchen Netzes kann deshalb erhebliche Schwierigkeiten und großen Aufwand erzeugen. Anstatt jeden Knoten einzeln und manuell auszubringen, sind in diesem Zusammenhang Massenprozesse vorgeschlagen worden, um diese Aufgabe zu vereinfachen und zu beschleunigen. Ein oft zitiertes Beispiel ist die Idee, mit einem Hubschrauber oder Flugzeug den zu beobachtenden Bereich zu überfliegen, und die Sensorknoten einfach kontinuierlich abzuwerfen. Die Knoten verteilen sich zufällig am Boden und können ein Netz ausbilden, um ihren Messaufgaben nachzukommen. Die lokalen Messungen werden typischerweise lokal vorverarbeitet und z.B. zu einer Basisstation weitergeleitet, wo sie gesammelt werden. Um die Messwerte durch das Netz zu transportieren, werden Multihop-Protokolle eingesetzt. In der wissenschaftlichen Literatur wurden hierzu bereits viele Mechanismen vorgeschlagen und miteinander verglichen [14]. Vereinfacht kann man sagen, dass ein Multihop-Protokoll eine Datenroute in einem Sensornetzwerk sucht, auf der eine Nachricht von Knoten zu Knoten bis zum Ziel weitergeleitet werden kann. Welche Routen in einem solchen Zusammenhang bevorzugt werden, kann nicht eindeutig beantwortet werden. Es kann die kürzeste, energiesparendste, schnellste, usw. Route sein. Damit der Datentransport von einer Quelle zu einem Ziel durchgeführt werden kann, muss jedoch mindestens eine solche Route existieren.

Wenn der Installationsprozess — wie oben erwähnt — zufällige Verteilungen von Sensorknoten erzeugt, so kann es zu Schwierigkeiten kommen, wenn die Knoten beginnen, Datenrouten aufzubauen: durch die zufällige Positionierung kann nicht garantiert werden, dass eine Multihop-Datenroute von allen Knoten zur Basisstation gefunden werden kann. Abhängig von der Verteilung (der Dichte) und Sendeleistung der Knoten, kann es passieren, dass einige Knoten als Gruppe separiert werden und keine Verbindung mehr zu den restlichen Sensorknoten oder zur Basisstation aufbauen können; sie sind vom restlichen Netz abgeschnitten. Es kommt zu einer sogenannten *Partitionierung*.

Dies kann nur über geplante Reserven und Redundanzen im Netz vermieden werden. Entweder die Knotendichte oder die Funkreichweite müssen deutlich erhöht werden, um eine insgesamt gute Konnektivität zwischen den Knoten zu erreichen.

6.1 Schlechte Abdeckung

Der oben genannte zufällige Installationsprozess ist nur ein Beispiel, wie eine insgesamt schlechte Konnektivität entstehen kann. Denn nicht nur der Installationsprozess, sondern auch andere Einflüsse wie z.B. *Mobilität* von Knoten kann zu unvorteilhaften Netzwerk-Topologien führen. Isolation und Gruppenbildung von Knoten kann viele Gründe haben, die unvermeidbar sind, und auch nicht von Administratoren beeinflusst werden können. Die nun folgenden Abschnitte erklären einige der wichtigsten Gründe, warum Verbindungen in Sensornetzwerken abbrechen und das Netz partitionieren können:

Zufälliger Installationsprozess: Wenn ein Netzwerk durch einen zufälligen Verteilungsprozess (wie z.B. Abwerfen aus einem Flugzeug) installiert wird, kann die Verteilung der Knoten natürlich nicht genau vorhergesagt werden. Wenn es sich um eine sehr raue Umgebung mit Felsen und Hügeln oder sogar Wald handelt, kann es zu einer Vielzahl von nachteilhaften Positionen kommen: Einige Knoten können bei der Landung bereits zerstört werden, andere können in Flüsse fallen und davontreiben. Es ist auch denkbar, dass Knoten in Niederungen landen, die von abschirmenden Felsen und Gewächsen umgeben sind, so dass keine Funkkommunikation nach außen möglich ist. All diese beispielhaften Einflüsse führen zu einer ungleichen Verteilung und Ausbildung des Netzes — typischerweise mit Gruppenbildung. Es wird auch zu einem Ungleichgewicht in der Bedeutung von Knoten kommen: Einige sind redundant, da sie direkt neben einem anderen Knoten gelandet sind, andere haben essentielle Bedeutung, da sie die einzige Kommunikationsbrücke zwischen zwei großen Gruppen von Knoten bilden.

Veränderungen in der Umgebung: Wenn Sensorknoten in einem dynamischen Umfeld ausgebracht werden, wird sich die Konnektivität ständig verändern. Um mehr Verständnis über ein solches Szenario zu sammeln, wurde im Zusammenhang mit dieser Arbeit ein Interview mit Verantwortlichen des LOFAR [1] Projekts durchgeführt. Diese hatten über mehrere Monate ein Sensornetzwerk in einem Kartoffelfeld ausgebracht [77]. Die in dem Zusammenhang von Konnektivität und Topologien interessanteste Erfahrung war, dass Funkverbindungen sehr sensitiv auf Regen und Gewächs sind. Am Anfang des Experiments waren die Kartoffelpflanzen sehr klein und die verteilten Sensorknoten hatten gute Verbindungen untereinander. Mit dem Wachstum der Pflanzen verschlechterten sich diese Verbindungen zuneh-

mend und führten an zufälligen Stellen zu wochenlangen Unterbrechungen der Funkverbindungen.

Wear-out: Wenn Sensorknoten für Messaufgaben über lange Laufzeiten ausgebracht werden, so verfügen sie für gewöhnlich über eine unabhängige Energiequelle: eine Batterie. Die Lebensdauer eines Sensorknotens ist dann auf die Lebensdauer dieser Batterie beschränkt. Am Ende der geplanten Laufzeit werden die Funkknoten anfangen „auszusterben“. Wichtige Verbindungsstellen werden aufhören zu arbeiten und möglicherweise Gruppen von Sensoren von anderen abtrennen. Hier ist es wahrscheinlich, dass es zu Partitionierungen des Netzwerkes kommt.

Mobilität von Knoten: Dieser Fall ist ein sehr wichtiger Grund für eine Partitionierung des Netzwerkes. Wenn Knoten in einem Netz mobil sind und nicht durch stationäre Infrastruktur unterstützt werden, kann die Konnektivität stark schwanken. Darüber hinaus werden die Veränderungen auch besonders schnell auftreten. Solche mobilen Szenarien sind z.B. die Beobachtung von Herden von Tieren mit Sensorknoten, die Beobachtung von Tieren, die in Flüssen oder im Meer schwimmen, Beobachtung von Eisschollenbewegungen oder ähnliche Effekte natürlicher Bewegung.

6.2 Das Konzept von kooperativem Senden

Kooperatives Senden ist ein ideales Mittel, um den Effekten von schlechter Konnektivität entgegenzuwirken. Mit kooperativem Senden kann eine Gruppe von Sensoren ihre Sendeleistung kombinieren, und so eine insgesamt höhere Gesamtleistung und damit eine größere Reichweite erreichen. Um dies zu tun, müssen Knoten gemeinsam identische Signale aussenden. Je mehr Knoten kooperativ senden, desto größer wird die Sendeleistung sein. So kann eine Gruppe von Knoten ein sehr weit entferntes Ziel erreichen. In Abbildung 6.1 sind zwei Funkknoten mit ihren einzelnen Reichweiten und ihrer gemeinsamen Reichweite dargestellt.

In der Literatur taucht dieses Konzept von kooperativem Senden zur Erhöhung der Konnektivität in schwach besetzten Szenarien nur selten auf. In [75] und [71] wird ein System diskutiert, das auf Rundruf mit kooperativem Senden aufgebaut ist. Zur Analyse werden die Knoten als eine kontinuierliche Dichte modelliert. Dies vereinfacht die Modellierung erheblich. Die vorliegende Arbeit befasst sich stattdessen insbesondere mit inhomogenen Verteilungen, die durch Effekte entstehen, wie sie in Abschnitt 6.1 dargestellt wurden. Die Effekte durch

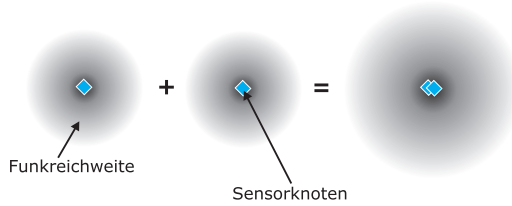


Abbildung 6.1. Erhöhung der Funkreichweite durch Summation der Sendeleistungen

„Aussterben“ von kritischen Relaisstationen, heterogene und dynamische Umgebungen sind große Herausforderungen für den Erfolg von Sensornetzwerken. Es soll nur die Thematik von kooperativem Senden — also überlagerter Funksignale — zur Erhöhung der Konnektivität in Sensornetzwerken detailliert diskutiert werden. Auf den Energieverbrauch soll in diesem Zusammenhang nicht eingegangen werden. Der interessierte Leser sei an dieser Stelle auf eine Arbeit von Yao-Win Hong [26] verwiesen.

6.2.1 Ausbreitungs- und Kanalmodell

Es sei P_{tx} die nominale Sendeleistung eines Knotens. Es wird weiter angenommen, dass die Sendeleistung aller Knoten gleich ist. Mit $P_{rx,j \leftarrow i}$ wird die Empfangsleistung bezeichnet, die Knoten i in Knoten j beim Senden erzeugt. Eine Empfangsleistung $P_{rx,j \leftarrow i}$ über einer festgelegten Schwelle P_{th} soll genügend Rauschabstand (SNR) erzeugen, dass der Empfänger die Nachricht dekodieren kann.

Für das Kanalmodell zwischen Knoten i und j werden drei Einflüsse auf das Funkssystem berücksichtigt:

Wegverlust: Die Dämpfung aufgrund von zunehmender Entfernung wird als radiale Dämpfung $\sim 1/r_{i,j}^\beta$ modelliert. Dabei ist r der euklidische Abstand und β der Dämpfungsexponent. Dieser wird in Modellierungen typischerweise zwischen 2 und 4 gesetzt. Die mittlere maximale Empfangsentfernung kann jetzt angegeben werden: $r_{th}^\beta = \frac{P_{tx}}{P_{th}}$

Antenneneinfluss: Antennen von drahtlosen Sensorknoten sind typischerweise gedruckte Antennen oder einfache Spulen. Es kommen auch häufig Strip-, Dipol- oder Wurfantennen zum Einsatz. Keine dieser Antennen hat ein omnidirektionales Verhalten. Deshalb wird ein zusätzlicher, zufälliger Einflussfaktor $\alpha_{i,j} \sim \mathcal{N}(1, 0.5)$ zwischen zwei Knoten angenommen, der die zufällige Orientierung der Antenne widerspiegelt.

Kanalmodell: Wegen erwarteter Mehrwegeausbreitung wird hier ein Rayleighprozess angenommen.

Diese Faktoren bilden zusammen die Basis für die Simulationen in den folgenden Abschnitten. Es wird weiter angenommen, dass das System die kooperativen Sendungen voll nutzen kann, also sich der Leistungsgewinn im Empfänger voll auswirkt. Eine Gruppe G von Knoten, die alle miteinander in Verbindung stehen, können ihre Sendeleistung bezüglich eines Empfängers j akkumulieren:

$$P_{rx,j \leftarrow G} = \sum_{i \in G} P_{rx,j \leftarrow i} \quad (6.1)$$

Mit den genannten Einflüssen entsteht:

$$P_{rx,j \leftarrow G} = \sum_{i \in G} \alpha_{i,j} P_{tx} \left(\frac{r_{th}}{r_{i,j}} \right)^\beta \quad (6.2)$$

6.2.2 Transportszenarien

Es wird nun zwischen den zwei wichtigsten Transportszenarien für Sensornetzwerke unterschieden: Das *Peer-to-Peer*-Szenario und das *Basisstations*-Szenario. In Letzterem wird angenommen, dass es eine Basisstation in der Mitte des Feldes gibt, die eine hohe Sendeleistung hat und so die Sensorknoten direkt anfunken kann. Die Sensorknoten hingegen verfügen nur über sehr limitierte Reichweiten und können nicht alle direkt die Basisstation erreichen. Im Basisstationsszenario findet sämtliche Kommunikation ausschließlich zwischen Knoten und Basisstation statt. Kommunikation zwischen Knoten gibt es nur zur Weiterleitung von Nachrichten von oder zu der Basisstation. Im Peer-to-Peer-Szenario sollen Nachrichten zwischen zufällig gewählten Knoten ausgetauscht werden. Solche Szenarien sind interessant, wenn z.B. Messwerte direkt innerhalb des Netzwerkes ausgewertet werden und gar keine Basisstation vorhanden ist. Im weiteren Verlauf

werden Knoten als *verbunden* bezeichnet, wenn sie in der Lage sind, bidirektional Informationen auszutauschen. Hierzu können sie Multihop- oder kooperatives Senden verwenden.

Für alle Szenarien wird weiterhin eine sich ständig ändernde Umgebung angenommen. Die Gründe hierfür sind in Abschnitt 6.1 erläutert. Für schnell ändernde und unvorhersagbare Umgebungen gibt es eine einfache Lösung für den Datenaustausch: Da sich die Wege zu den Kommunikationspartnern ständig ändern können, wird die Kommunikation ausschließlich auf *Rundruf* basiert. Für alle Multihop-Szenarien impliziert dies die Verwendung von *flooding*.

6.2.3 Prinzipien der kooperativen Kommunikation

Um kooperatives Senden mit herkömmlichen Verfahren vergleichen zu können, ist es notwendig, klar zwischen den verschiedenen Ansätzen (Prinzipien) der Kommunikation zu unterscheiden. Betrachtet man schwach besetzte Topologien, so ist es klar, dass Multihop-Kommunikation an seine Grenzen stößt, wenn die geringe Knotendichte zu Partitionierungen und Gruppenbildungen im Netz führt. In solchen Fällen kann jedoch kooperative Kommunikation helfen, die zerstörten Verbindungen zu „reparieren“. Wie nun genau Vorteil aus kooperativem Senden gezogen werden kann, ist zunächst unklar. Die optimale Leistungssteuerung für kooperative Kommunikation z.B. wurde als NP-vollständiges Problem erkannt [26]. Deshalb sollen für die Kommunikation pragmatische Ansätze vorgestellt werden, die nun die vier Prinzipien der Kommunikation darstellen:

1. Herkömmliches Multihop-(Fluten)
2. Kooperatives Senden als Wellenausbreitung
3. Akkumulierendes kooperatives Senden
4. Ideales Hybridverfahren zwischen Multihop- und akkumulierendem, kooperativem Senden

In den folgenden Abschnitten wird die Situation diskutiert, dass ein Knoten (Nummer 0) eine Nachricht an ein Ziel (Nummer 6) weiterleiten will. An diesem Beispiel werden die vier Kommunikationsprinzipien erklärt.

Abbildung 6.2 zeigt das Referenzszenario. Das kleine Netzwerk in diesem Beispiel ist stark partitioniert. Knoten 0 kann nur mit Knoten 1 kommunizieren; es gibt ein weiteres Cluster, bestehend aus Knoten 3, 4 und 5. Knoten 2 und 6 sind

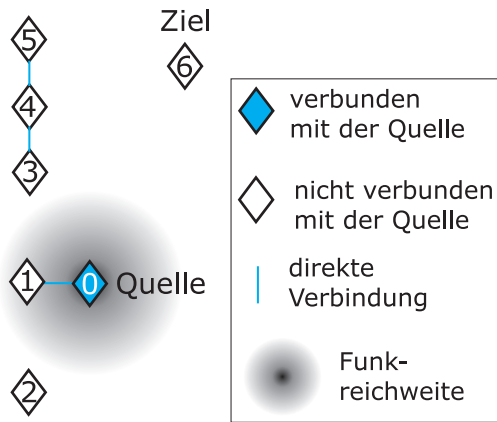


Abbildung 6.2. Das Kommunikationsszenario: Knoten 0 (Quelle) will eine Nachricht an Knoten 6 (Ziel) verschicken

vollständig isoliert. Der Transportprozess hat keine weitere Kenntnis über die Topologie, die Knoten könnten mobil sein und die Umgebung dynamisch. Deshalb wird — wie schon erwähnt — *Rundruf*- Kommunikation eingesetzt werden.

Herkömmliche Multihop-Kommunikation

Wie in Abbildung 6.2 zu sehen, kann Knoten 0 mit Knoten 1 kommunizieren. Knoten 1 kann die Nachricht von Knoten 0 allerdings an keinen weiteren Kommunikationspartner ausliefern. Somit ist der Transport der Nachricht von Knoten 0 schon im nächsten Schritt bei Knoten 1 zu Ende. Die Nachricht kann nicht an das Ziel (Knoten 6) ausgeliefert werden, da die Distanzen zwischen den Knoten zu groß sind.

Kooperatives Senden als Wellenausbreitung

Wenn Knoten dieses Verfahren einsetzen, dann werden sie jede Nachricht einmal wiederholen. Sie werden dies zusammen mit all denjenigen Knoten tun, die

ebenfalls direkt zuvor diese Nachricht empfangen haben. Dieses Kommunikationsprinzip entspricht dem Prinzip von *opportunistic large arrays* in [71]. Die Nachricht wird wie eine Wellenfront durch das Netz propagiert. Für unser Szenario heißt dies, dass nach der Aussendung von Knoten 0, alle Knoten, die diese Nachricht empfangen haben, gemeinsam (also kooperativ) diese Nachricht wiederholen werden. Unglücklicherweise ist nur Knoten 1 Empfänger der Nachricht, und so scheitert auch dieses Verfahren im zweiten Schritt. Knoten 1 kann keinen weiteren Knoten erreichen und die Welle stirbt aus.

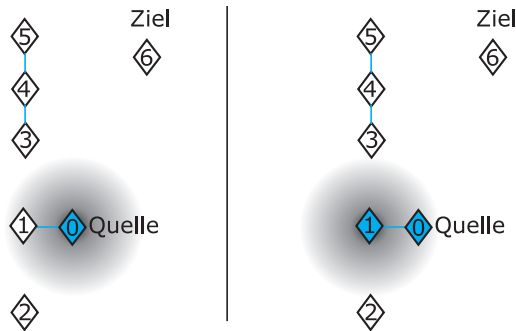


Abbildung 6.3. Das Kommunikationsszenario mit kooperativem Senden und Wellenausbreitung

Akkumulierendes kooperatives Senden

Dieses Verfahren entsteht aus einer leichten Modifikation des vorhergehenden Verfahrens. Es ist ähnlich zu dem *cumulative increment algorithm* in [26]. Knoten, die eine Nachricht empfangen, werden diese nicht nur einmal, sondern mehrere Male wiederholen. Die Anzahl der Wiederholungen wird als Systemparameter festgesetzt. Mit diesem Verfahren sind nun in dem Referenzbeispiel die ersten Erfolge zu verzeichnen: Abbildung 6.4 zeigt die Situation der ersten zwei Schritte. Nachdem Knoten 0 die Nachricht an Knoten 1 ausgeliefert hat, wiederholen sie nun *beide* diese Nachricht und akkumulieren so ihre Sendeleistung (linke Seite in Abbildung 6.4). So können die nächsten zwei Knoten (2 und 3) erreicht werden. Im darauffolgenden Schritt umfasst die Gruppe der kooperativ Sendenden bereits

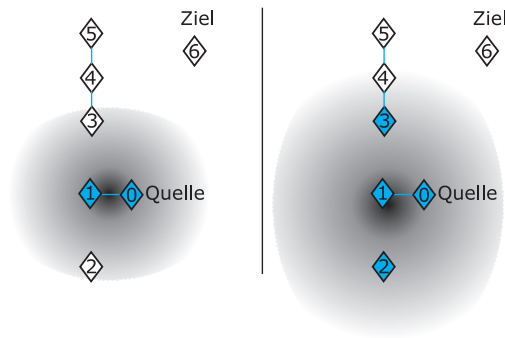


Abbildung 6.4. Das Kommunikationsszenario und akkumulierendes, kooperatives Senden

die Knoten 0, 1, 2 und 3. Abbildung 6.5 zeigt nun noch die letzten zwei Schritte, in denen Knoten 4 und 5 auch noch in die Gruppe der Sendenden akkumuliert werden. Die akkumulierte Sendeleistung aller reicht nun aus, um Knoten 6 zu erreichen. Für die Auslieferung der Nachricht ist es nicht nötig, weitere Kommunikationsparameter festzulegen oder Routen zu finden. Die Knoten wiederholen einfach eine empfangene Nachricht einige Male und es ist kein weiterer Aufwand nötig.

Hybride Multihop- und kooperative Kommunikation

Die Idee dieses Prinzips ist es, Multihop-Kommunikation zu verwenden, wenn es möglich ist, und kooperative Kommunikation zu verwenden, wann immer es nötig ist. Abhängig von der vorliegenden Topologie, kann die Entscheidung, ob nun Multihop- oder kooperative Kommunikation im nächsten Schritt verwendet werden soll, nicht einfach gefunden werden. Wenn der nächste Schritt kooperative Kommunikation ist, ist es auch schwer zu entscheiden, welche Knoten der mit Multihop verbundenen Gruppe nun daran teilnehmen sollten. Diese Entscheidungen können nur mit Wissen über die Topologie sinnvoll getroffen werden. Deshalb ist auch hier ein pragmatischer Ansatz notwendig. Es wird eine *alternierende* Kommunikation von kooperativem und klassischem Multihop eingesetzt. Nach *einer* kooperativem Kommunikation werden alle neuen Knoten eine ge-

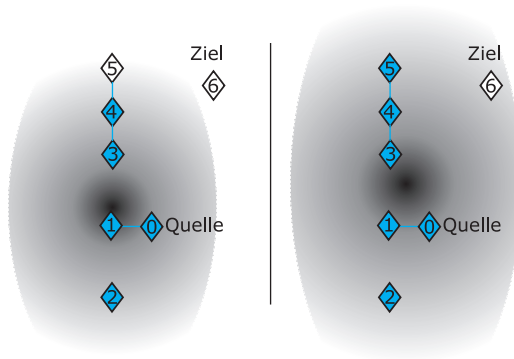


Abbildung 6.5. Das Kommunikationsszenario mit akkumulierendem, kooperativem Senden

wisse Zeit versuchen, mit Multihop-Kommunikation (also flooding) noch weitere Partner zu akquirieren. Anschließend wird wieder ein kooperatives Senden mit der alten Gruppe und den neu dazu gewonnenen Kommunikationspartnern durchgeführt. Ob ein Knoten sich selbst als „neuer“ Partner für einen Kommunikationsprozess einstuft, kann durch die Sequenznummer der zu übermittelnden Nachricht festgestellt werden. Die nötige Zeit, die ein neuer Knoten braucht, um alle erdenklichen Partner über Multihop zu akquirieren, ist unbekannt und kann nur durch eine sinnvolle Schätzung festgelegt werden. Diese Zeit wird also als ein Systemparameter, eine Deadline, verstanden. Die Knoten würden z.B. kooperativ senden, die alten und neu erreichten Knoten hätten dann wieder eine gewisse Zeit, weitere Partner durch Multihop zu akquirieren, bis sie dann am Ende dieser Deadline *alle* gemeinsam kooperativ senden. Die Zahl der Iterationen dieses Prozesses ist wiederum ein Systemparameter. Die Abbildungen 6.6 und 6.7 erklären dies im Detail. Mit der ersten Runde Multihop-Rundrufe kann Knoten 0 den Knoten 1 finden. Mit diesem als Partner kann Knoten 0 dann kooperativ senden. Dies ist auf der linken Seite in Abbildung 6.6 zu sehen. Nach diesen ersten zwei Schritten werden Knoten 2 und 3 erreicht und können als neue Knoten zunächst die Nachricht auch mit Multihop weiterleiten. Diese Schritte sind auf der rechten

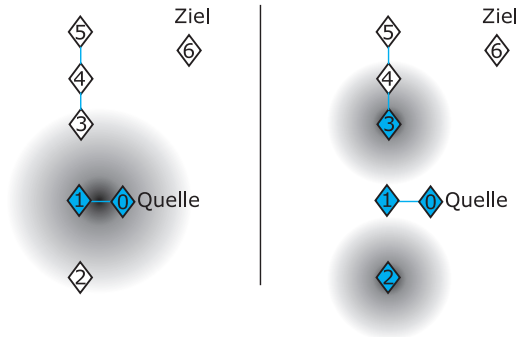


Abbildung 6.6. Das Kommunikationsszenario mit dem *hybriden* Verfahren

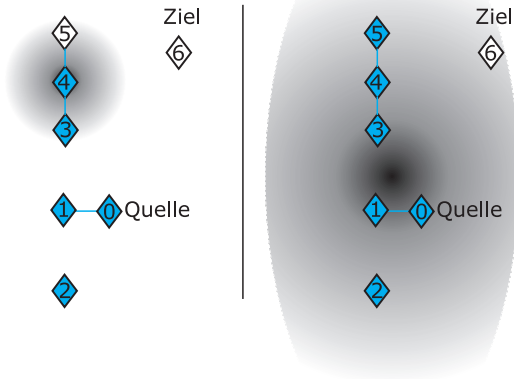


Abbildung 6.7. Das Kommunikationsszenario mit dem *hybriden* Verfahren

Seite von Abbildung 6.6 und auf der linken Seite von Abbildung 6.7 zu sehen. Anschließend wird der alternierende Prozess fortgeführt und alle neuen und alten Partner senden wieder kooperativ. Dies ist in Abbildung 6.7 auf der rechten Seite zu sehen. Dieser wechselnde Prozess zwischen Multihop und kooperativem Senden wird fortgeführt, bis eine vorbestimmte Anzahl von Wiederholungen erreicht ist.

6.3 Erhöhung der Reichweite

Sensorknoten können ihre Reichweite erhöhen, indem sie kooperativ mit anderen Knoten senden. Mit der erhöhten Reichweite ist es offensichtlich, dass die gesamte Abdeckung und Erreichbarkeit verbessert wird. Isolierte Knoten und Cluster haben so eine völlig neue Möglichkeit, wieder Verbindung zum restlichen Netz aufzunehmen. Zur weiteren Analyse und Illustration der Verbesserungen wurden Simulationen durchgeführt, in denen zufällige Szenarien betrachtet wurden. In den Simulationen werden die Knoten gleichverteilt in der Gesamtfläche positioniert. Für jede gegebene Knotenanzahl wurden 100 verschiedene, zufällige geometrische Konstellationen (Topologien) simuliert, auf denen jeweils die verschiedenen Protokolle getestet wurden. In Tabelle 6.1 werden die technischen Parameter der Simulation zusammengefasst. In der Simulation wurden verschiedenen Kommunikationsprinzipien unterschieden. Diese werden nun anhand der Benennung aus Abbildungen 6.8 und 6.9 erklärt:

- „multihop“: dies ist das normale Multihop-Prinzip, wie es aus herkömmlichen Netzen bekannt ist. Knoten leiten Pakete Stück für Stück weiter. Hier wurde als Protokoll *flooding* verwendet.
- „acc.co.tr.1“: dies entspricht dem *akkumulierenden kooperativem Senden*. Der Startknoten sendet ein Paket aus; alle Empfänger wiederholen das Paket einmal zusammen mit dem Startknoten. Danach terminiert dieser Prozess.
- „acc.co.tr.2“: dies entspricht dem Verfahren „acc.coo.tr.1“, es wird jedoch nach dem Wiederholungsschritt noch ein weiterer Wiederholungsschritt mit allen Knoten durchgeführt, die bis dahin das Paket empfangen konnten.
- „acc.co.tr.3“: entsprechend den zwei vorangehenden Verfahren, wird nun noch ein Wiederholungsschritt hinzugeführt.

- „hybrid 1“: In diesem Verfahren sendet der Startknoten die zu versendende Nachricht an alle Partner in seiner Multihop-Umgebung. Zusammen mit diesen Partnern wird dann ein kooperatives Senden durchgeführt.
- „hybrid 2“: Dieses Verfahren ist wiederum die schrittweise Erweiterung des hybrid 1-Verfahrens. Die mit hybrid 1 erreichten Knoten informieren alle ihre Multihop-Partner und anschließend wird kooperativ gesendet mit allen hybrid 1 und allen neuen Partnern.

Simulationsfläche	500m x 500m = 250.000m ²
Anzahl der Knoten	10..200
Mittlere nominale Funkreichweite	50m
Fading-Exponent	$\beta = 2$
Anzahl der Topologien pro Knotendichte	100

Tabelle 6.1. Simulationsparameter

Die relativen Anzahlen in den Graphiken 6.8 bis 6.11 sind Mittelwerte über die jeweils 100 Topologien pro gegebener Anzahl von Knoten.

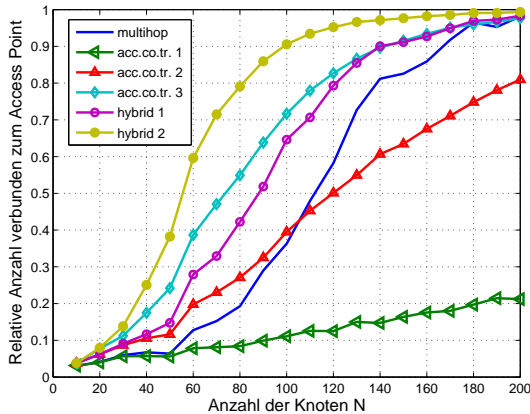


Abbildung 6.8. Konnektivität der Knoten mit ihrer Basisstation, Fläche=500m x 500m, Funkreichweite \approx 50m

In Abbildung 6.8 ist die mittlere Anzahl von Knoten dargestellt, die Verbindung zur Basisstation hat. Eine Verbindung zur Basisstation bedeutet hier, dass ein Knoten in der Lage ist, eine Nachricht *zur* Basisstation zu leiten. Die Rückrichtung wird hierbei nicht betrachtet, da davon ausgegangen wird, dass die Basisstation über hohe Sendeleistung verfügt und so die Knoten direkt erreichen kann. Zunächst lässt sich feststellen, dass für sehr kleine Knotenanzahlen alle Protokolle gleich „schlecht“ arbeiten. Das liegt daran, dass bei sehr geringer Knotendichte der Effekt des kooperativen Sendens nicht genutzt werden kann. Die Knoten sind alle isoliert und finden keine Partner, mit denen sie gemeinsam senden könnten. Genauso ist es auch klar, dass für große Knotendichten alle Protokolle, die Multihop unterstützen, eine 100%-ige Abdeckung oder Konnektivität erzeugen. Protokolle wie das akkumulierende Senden (z.B. „acc.co.tr.1“ in Abbildung 6.9) werden bei nur einer Wiederholungstiefe nicht 100%-ige Abdeckung erreichen, egal wie viele Knoten absolut anwesend sind: es wird eine große Anzahl von Knoten geben, die nicht in Zweisprungentfernung an den jeweiligen Kommunikationspartner verbunden werden können.

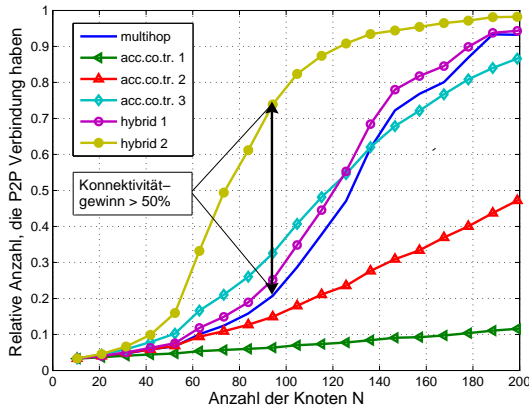


Abbildung 6.9. Peer-to-Peer-Konnektivität von Knoten

In Abbildung 6.9 werden die Verbindungen zwischen beliebigen Knotenpaaren betrachtet. Eine Verbindung gilt dann als etabliert, wenn bidirektionales Austauschen von Paketen zwischen zwei Knoten (mit dem jeweiligen Kommunikati-

onsprinzip) möglich ist. In Abbildung 6.9 ist ein sehr interessantes Maß eingezeichnet. Es ist der Konnektivitätsgewinn mit kooperativem Senden. Wenn etwa 90 Knoten anwesend sind, die mit klassischen Multihop-Verfahren kommunizieren, sind nur 20% der Verbindungen zwischen den Knoten möglich. Dieser Wert kann mit keinem Routingverfahren verbessert werden. Er ist durch die physikalische Funkreichweite limitiert. Durch den Einsatz von kooperativem Senden kann — unter sonst gleichen Voraussetzungen — die Konnektivität auf über 70% erhöht werden. Diese Verbesserung von über 50% in der gesamten Konnektivität im Netzwerk ist ein eindrucksvolles Beispiel für die Leistungsfähigkeit von kooperativem Senden und den dafür notwendigen überlagerten Funksignalen.

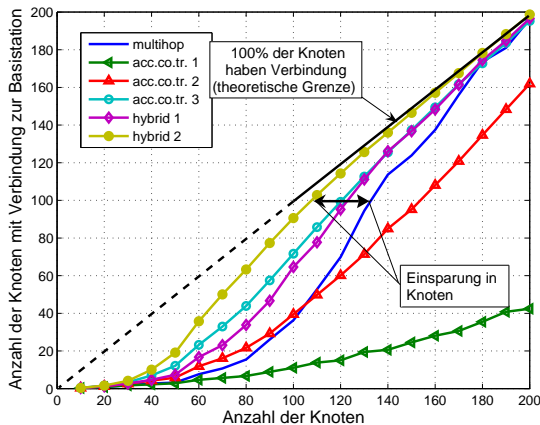


Abbildung 6.10. Anzahl der Knoten mit Verbindung zur Basisstation

Das nächste interessante Maß ist die Anzahl der Knoten, die tatsächlich nötig sind, um daraus eine gewisse Anzahl von Knoten zu erhalten, die Verbindung zur Basisstation haben. Es geht hier also um die notwendige Redundanz. Wenn Knoten zufällig in einem Bereich ausgebracht werden, so werden nicht alle Knoten Verbindung zur Basisstation haben. Wenn jedoch eine gewisse Anzahl von Knoten eine garantierte Verbindung haben soll, so ist es notwendig, mehr als diese minimale Anzahl auszubringen. In Abbildung 6.10 kann man die Anzahl der eingesetzten Knoten mit der Anzahl der Knoten vergleichen, die tatsächlich eine Verbindung zur Basisstation herstellen konnten. Die theoretische Grenze der

Konnektivität ist dann erreicht, wenn alle Knoten zur Basisstation verbunden sind. Die theoretische Grenze ist in Abbildung 6.10 als Winkelhalbierende eingezeichnet und ist für die Anzahlen unter 100 Knoten gestrichelt, da für die gegebene Fläche selbst bei idealer geometrischer Konstellation mindestens 100 Knoten für eine Vollabdeckung notwendig sind. In Abbildung 6.10 sieht man jedoch, dass die zufällige Verteilung und die zufälligen Variationen in der Funkverbindung dazu führen, dass von 100 Knoten nur 40 tatsächlich eine Verbindung zur Basisstation etablieren können. Es ist also klar, dass mehr als 100 Knoten nötig sein werden, um eine Verbindung von 100 Knoten zur Basisstation zu erreichen. Aus Abbildung 6.10 kann man für normalen Multihop-Betrieb 130 nötige Knoten erkennen, wenn 100 Knoten tatsächlich Verbindung zur Basisstation haben sollen. Wenn stattdessen das „hybrid 2“-Protokoll verwendet wird, sinkt diese Zahl auf 110. Mit der gleichen Abdeckung und den sonst gleichen Parametern für die Funkreichweite und dem Installationsprozess ist also mit dem hybriden Protokoll weniger Redundanz nötig. Dies reduziert die Kosten, wenn ein solches Netzwerk installiert wird.

Insgesamt werden die besten Ergebnisse von den hybriden Protokollen erreicht. Dies ist sofort klar: die hybriden Protokolle vereinigen die Vorteile von Multihop- und kooperativem Senden und können so die größten Reichweiten und damit die beste Konnektivität erzeugen. Andererseits sind die hybriden Protokolle auch am Aufwendigsten in der Implementierung. *Akkumulierendes, kooperatives Senden* kann nicht effektiv arbeiten, wenn nur wenige Wiederholungsschritte eingesetzt werden. Die schlechten Ergebnisse (besonders für den Fall „acc.co.tr. 1“) sind darauf zurückzuführen, dass die Knoten nur ihre direkten Nachbarn zur Hilfe nehmen und die Multihop-Möglichkeiten nicht ausschöpfen. Trotzdem kann normaler Multihop-Verkehr schon bei nur drei Stufen übertroffen werden. In Abbildung 6.8 ist dies zu sehen: Mit „acc. co. tr. 3“ ist eine höhere Konnektivität als bei „Multihop“ zu erreichen.

6.3.1 Grenzverhalten

Nachdem nun anhand von Simulationen von ausgewählten Szenarien und Protokollen gezeigt wurde, welche Beiträge kooperatives Senden für Sensornetzwerke liefern kann, sollen nun noch die maximalen Gewinne bestimmt werden, die mit kooperativem Senden überhaupt erreicht werden können. Hierfür wurden wieder zufällige Topologien simuliert und die Protokolle „akkumulierendes kooperatives

Senden “ und „hybrides kooperatives Senden“ betrachtet. Diese Verfahren wurden nun in verschiedenen Wiederholungstiefen betrachtet, um festzustellen, welchen Beitrag ein weiterer Wiederholungsschritt noch liefern kann. In Abbildung 6.11 kann man die Konnektivitäten für verschiedene Protokolle und verschiedene Wiederholungstiefen (level) erkennen.

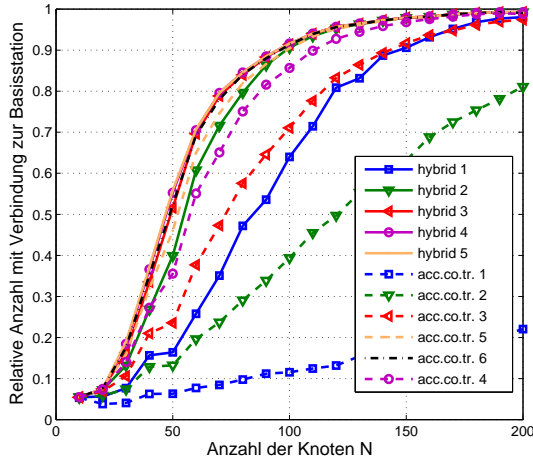


Abbildung 6.11. Grenzgewinn mit kooperativem Senden

Die wichtigste Feststellung, die aus den Simulationen gewonnen werden kann, ist zunächst, dass beide Protokolle (akkumulierendes und hybrides kooperatives Senden) den gleichen mittleren Konnektivitätsgewinn erzielen können. Wenn also in einem System die maximal mögliche Konnektivität erreicht werden muss, so ist es letztlich nicht abhängig vom eingesetzten Protokoll. In dem hier simulierten Szenario erreicht das hybride Verfahren den maximalen Gewinn schon in der dritten Wiederholungstiefe. Dies bedeutet, dass nach drei Runden des hybriden Prozesses von kooperativem Senden und Multihop-Kommunikation, die maximale Konnektivität erreicht wird. Mehr Wiederholungen können kaum weitere Verbesserungen erreichen. Für das akkumulierende Protokoll wird eine vergleichbare Konnektivität erst nach sechs Wiederholungen erreicht. Das hybride Protokoll scheint also zunächst der klare Gewinner; dies ist jedoch nur teilweise richtig.

Das hybride Protokoll schließt auch Multihop-Kommunikation in seinen Prozess mit ein. Also ist das hybride Protokoll „komplexer“ und erfordert mehr Koordination und Kommunikation in jedem Schritt. Der Vergleich der beiden Protokolle müsste hier auch den Implementierungsaufwand oder den Energieverbrauch mit berücksichtigen.

6.3.2 Energieverbrauch und Implementierungsaufwand

Wie bereits in Abschnitt 6.2.2 erwähnt, werden für sämtliche Kommunikation verteilende Rundrufe verwendet. Das Ausbreitungsmodell hierzu wurde in Abschnitt 6.2.1 vorgestellt. Da diese Arbeit nicht im Fokus hat, den Energieverbrauch solcher Anwendungsprotokolle von überlagerten Funksignalen zu untersuchen, sollen hier nur einige prinzipielle Feststellungen gefasst werden: Die hybride Kommunikation — die die leistungsstärkste ist — hat auch den höchsten Energieverbrauch. Dies liegt an dem alternierenden Prozess von Multihop- und kooperativem Senden.

Es existiert ein prinzipieller Zielkonflikt zwischen Konnektivität und Energieverbrauch. Je mehr Wiederholungsschritte ein Protokoll durchläuft, desto mehr Energie wird verbraucht werden. Mehr Schritte erhöhen jedoch auch die Konnektivität. Für Topologien mit ausreichender Konnektivität kann man erwarten, dass Multihop-Lösungen energieeffizienter arbeiten werden. Für dünn besetzte Szenarien kann jedoch bei Multihop-Kommunikation mit dem Einsatz von mehr Energie keine weitere Verbesserung erzielt werden. Im Gegensatz dazu gilt für Protokolle, die kooperatives Senden einsetzen, dass der Einsatz von mehr Energie im Netz in den meisten Fällen zur Verbesserung der Konnektivität führt.

Der Implementierungsaufwand für die verschiedenen Protokolle ist sehr unterschiedlich. So finden alle Protokolle, die nur auf kooperatives Wiederholen bauen („acc.co.tr.x“), sehr einfache Implementierungen. Eine einkommende Nachricht wird schlicht ein- oder mehrmals hintereinander wiederholt und überlagert sich so mit den Wiederholungen anderer. Die hybriden Protokolle erfordern jedoch mehr Aufwand. Hier kommt zu den Wiederholungsschritten auch noch eine Multihop-Kommunikation. Diese muss nach einer vorgegebenen Zeit terminieren und möglichst effizient arbeiten.

6.4 Zusammenfassung

Mit überlagerten Funksignalen, die für kooperatives Senden eingesetzt werden können, kann die Konnektivität von Sensorknoten besonders in dünn besetzten Szenarien verbessert werden. In Abschnitt 6.1 wurden einige typische Fälle erläutert, wie es zu solchen dünn besetzten Topologien kommen kann, und weshalb sie nicht immer vermieden werden können. Kooperatives Senden kann speziell partitionierte Netzwerke mit separierten Gruppen von Knoten unterstützen. Eine separierte Gruppe kann mit kooperativem Senden ihre Reichweite erhöhen und so wieder an das restliche Netzwerk angeschlossen werden. *Dieser Effekt ist ein Beitrag von kooperativem Senden, der von Multihop-Protokollen nie erbracht werden kann.* Es können also mit diesem neuen Kommunikationsprinzip Probleme gelöst werden, die auf herkömmliche Weise nicht gelöst werden können. Zurückblickend auf Abbildung 6.11 kann man sagen, dass das hybride Verfahren die schnellste Konvergenz zum Optimum besitzt, es ist aber auch das komplexeste und am aufwendigsten zu implementieren. In Tabelle 6.2 sollen nun Anwendungsfelder und Aufwand der Verfahren verglichen werden. Für eine prakti-

Protokoll	Implementierungsaufwand	Anwendungsfelder
Multihop-Rundruf (flooding)	mittel	dichte Besetzungen statische Szenarien
Wellenausbreitung kooperatives Senden	niedrig	dichte Besetzungen Rundrufverkehr
akkumulierendes kooperatives Senden	niedrig	dünne Besetzungen statische, mobile Szenarien
hybrid kooperatives Senden	hoch	dünne Besetzungen statische Szenarien

Tabelle 6.2. Vergleich der Kommunikationsprinzipien

sche Implementierung scheint das akkumulierende kooperative Senden eine sehr interessante Option. Mit nur drei Wiederholungsschritten kann dieses Protokoll bereits die Leistungsfähigkeit von Multihop-Kommunikation erreichen mit ebenfalls vergleichbarem Energieaufwand. Aber das kooperative Senden hält darüber hinaus die Option vor, dass mit weiteren Wiederholungsschritten die Konnektivität weiter erhöht wird. Multihop-Kommunikation hingegen ist auf die direkte

Vermaschung der Funkknoten angewiesen. Die Implementierung von kooperativem Senden erzeugt kaum Overhead und ist sehr schlank. Für drahtlose Sensornetzwerke kann dieses kooperative Senden auch als Notfallverfahren eingesetzt werden, so dass nur in Fällen, in denen normale Multihop-Kommunikation scheitert, kooperatives Senden selektiv eingesetzt wird.

6.5 Bestehende Probleme für die Implementierung und weitere Verbesserungen

Mit den Kommunikationsverfahren, wie sie in diesem Abschnitt dargestellt wurden, wurde gezeigt, wie überlagerte Funksignale auf Protokollebene Verbesserungen in Netzwerken erzeugen können. Mit den Simulationen sind theoretische mittlere und maximale Grenzen für Gewinne etc. bestimmt worden. Die Protokolle müssen jedoch für einen echten Einsatz noch erheblich verbessert werden. So treten bei kooperativem Senden alle bekannten Probleme aus der Funkkommunikation auf. Beispiele sind hier *hidden* und *exposed terminals*. Diese Probleme müssen vor dem Hintergrund der überlagerten Funksignale neu diskutiert werden, da die Koordination einer ganzen Gruppe zusätzliche Schwierigkeiten und Aufwand erzeugt. Die Optimierung von Energieverbrauch ist auch ein wichtiger Teilaspekt, der in den Simulationen nicht betrachtet wurde. Hierfür soll nun das Beispielszenario aus Abbildung 6.12 betrachtet werden. Es sind zwei Gruppen von Sensorknoten zu sehen. Jede Gruppe ist in der Lage, mit normaler Multihop-Kommunikation innerhalb ihrer Gruppenteilnehmer zu kommunizieren. Die jeweils andere Gruppe kann jedoch mit einfacher Multihop-Kommunikation nicht erreicht werden, da die Entfernung zu weit ist. Die Protokolle, die in den letzten Abschnitten vorgestellt wurden, schlagen vor, mehrere Knoten zu einer Sendegruppe zu vereinen und durch kooperatives Senden die gesamte Reichweite zu erhöhen. Würde nun z.B. das Hybridverfahren auf die Gruppe 1 angewendet, so würde zunächst mit Multihop die gesamte Gruppe 1 sich auf eine Nachricht einigen, die dann gemeinsam versendet würde. Dabei wird keine Rücksicht auf die tatsächliche Konstellation genommen. Es wäre sicher effizienter, nur diejenigen Knoten senden zu lassen, die am nächsten an dem Ziel platziert sind. In Abbildung 6.12 sind diese Knoten mit einem weißen Kreis markiert. Die anderen Knoten würden nur einen unerheblichen Beitrag zur Sendeleistung beitragen, weil sie vom Ziel sehr weit entfernt sind. Nun ist es jedoch nicht a-priori bekannt, wo sich die nicht erreichte Gruppe 2 befindet. Die Mechanismen zur Optimierung

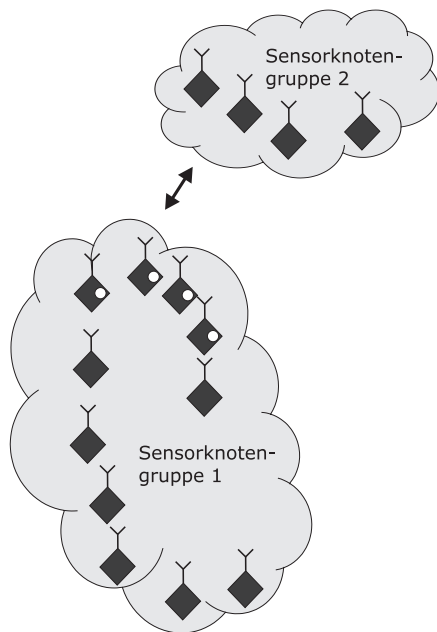


Abbildung 6.12. Zwei Gruppen von Sensorknoten, die nur über kooperatives Senden miteinander kommunizieren können.

wären hier recht komplex. So müsste z.B. zunächst eine Suche stattfinden und dann versucht werden, nur die nötigsten Funkknoten aus der gesamten Gruppe auszuwählen, um den gesamten Energieverbrauch zu verringern.

7 Cross-Layer Parameterschätzung mit überlagerten Funksignalen

In diesem Kapitel wird das Verfahren SDJS vorgestellt, das Datenübertragung basierend auf überlagerten Funksignalen ermöglicht. Mit diesem Verfahren kann man Daten von verteilten Quellen — also z.B. verstreuten Sensorknoten — schnell und effizient sammeln und verarbeiten. Dabei werden die Daten gleichzeitig im selben Funkkanal verschickt. Die dabei entstehenden Kollisionen werden von dem Verfahren berücksichtigt und führen nicht zu vollständigem Datenverlust, wie es sonst in klassischer Datenübertragung der Fall ist. Durch die spezielle Behandlung von Kollisionen ergibt sich jedoch eine grundlegende Eigenschaft dieses Verfahrens: es ist rein probabilistischer Natur. Das heißt, dass Daten, die mit diesem Verfahren kommuniziert werden, nicht in Paketform übertragen werden. Deshalb ist es auch nur für ganz spezielle Aufgaben — typischerweise Datenfusion in Sensornetzwerken — geeignet. Normale Datenübertragung ist damit nur schwer möglich. Es bietet stattdessen in anderen Bereichen interessante Merkmale: so ist z.B. die Kommunikationszeit, die nötig ist, verteilte Daten mit diesem Verfahren einzusammeln, unabhängig von der Anzahl der Teilnehmer. Es ist dadurch für Funknetze mit hoher Teilnehmeranzahl oder für mobile Szenarien besonders geeignet. Das Verfahren baut auf den Erkenntnissen der voranstehenden Kapitel auf und die Anforderung an die Hardware werden auch für dieses Verfahren genauso streng gehalten, wie für alle anderen Mechanismen, die in dieser Arbeit diskutiert werden. Neben den theoretischen Betrachtungen werden in diesem Kapitel auch verschiedene Anwendungen für SDJS aufgezeigt, wie z.B. der Einsatz für Warenwirtschaftssysteme im Einzelhandel.

Das Schichtenmodell von Kommunikationsprotokollen wird oft durch weitere Komponenten unterstützt, die dem typischen Schichtenaufbau widersprechen. Diese Komponenten sind nicht mehr eindeutig einer Schicht zugeordnet und übernehmen z.B. Aufgaben des Managements. Abbildung 7.1 aus [50] zeigt

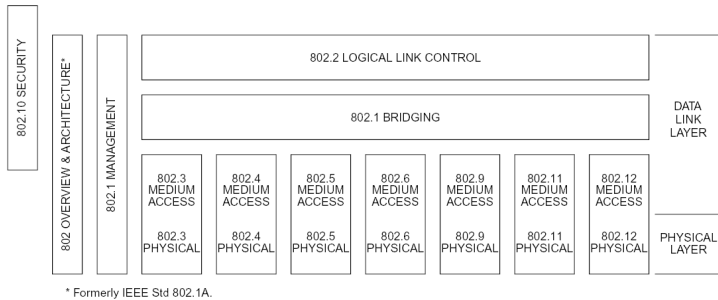


Abbildung 7.1. Architektur der IEEE 802.x Standards

die Schichtenarchitektur der 802.x IEEE-Standards. Auch hier ist ein Managementblock (802.1) vorgesehen, der Aufgaben der Systemlastverteilung, der Gesamtpformance und des Informationsaustausches zwischen den Schichten übernimmt. Die herausragende Aufgabe bei solchen schichtenübergreifenden Funktionsblöcken ist es, die Systeme zu stabilisieren und zu optimieren. Für einen Teilnehmer in einem Netz mit einem verteilten Zugriffsverfahren ist z.B. die Kenntnis der *Geräte mit gleichzeitigem Sendewunsch* eine wertvolle Information zur Optimierung der Kanalvergabe. Dies wurde in Kapitel 5.4 ausgiebig diskutiert. Weitere interessante Informationen, die in Managementmodulen verarbeitet werden, sind z.B. Informationen über

- die Geräte mit gleichzeitigem Sendewunsch (bzw. deren Anzahl),
- die aktiven Geräte (bzw. deren Anzahl),
- die Anzahl der wartenden Pakete,
- die Energieresourcen eines Teilnehmers.

Will man solche *Parameter* der Netzwerkschichten benutzen, um damit die Protokollabläufe zu optimieren, so müssen diese Parameter zunächst einmal bestimmt werden. Im einfachsten Fall sind diese Parameter Messwerte wie z.B. die Energiereserve, die über die Batteriespannung ermittelt werden kann. Schwieriger

zu behandeln sind Parameter, wie z.B. die *Anzahl der Geräte mit Sendewunsch*, die keine lokalen Messgrößen sind. Solche Parameter, die als verteilte Information vorliegen, müssen über Kommunikation zusammengetragen werden. Dies erfordert es einen erheblichen Aufwand an Energie und Bandbreite, der – besonders in drahtlosen, batteriebetriebenen Systemen – als kritisch zu beurteilen ist. Um z.B. die Anzahl der Geräte mit Sendewunsch zu bestimmen, müsste man in der klassischen Herangehensweise Rundrufanfragen an die Umgebung stellen und z.B. die Identifikationen der Teilnehmer auflisten, um somit die Gesamtanzahl zu bestimmen. Solche Prozesse werden im IP-Netzwerk z.B. durch den Befehl PING ermöglicht. Dieser Prozess ist jedoch in vielen Fällen nicht geeignet um eine ausreichend schnelle Reaktionszeit zu erzeugen. Nach der Rundrufanfrage kommt es zu einer Antwortenexplosion aller Teilnehmer, was durch erhöhte Kollisionen eine erhebliche Zeit dauert. Für drahtlose Anwendung ist es besonders kritisch: die Teilnehmer sind typischerweise mobil und so passiert es, dass der zu bestimmende Parameter schon gar nicht mehr gültig ist, bis sämtliche Antworten eingesammelt wurden.

Dieses Kapitel befasst sich nun mit der Problemstellung, wie eine Parameterschätzung in drahtlosen Systemen beschleunigt werden kann. Die zu betrachtenden Informationen liegen stets in verteilter Form vor. Hierzu kommen überlagerte Funksignale zum Einsatz.

Das in den folgenden Abschnitten dargelegte Verfahren ermöglicht eine sehr schnelle Parameterschätzung und basiert auf überlagerten Funksignalen. Es verwendet eine synchrone, verteilte Signalisierung und wurde deshalb *Synchronous Distributed Jam Signalling*, kurz *SDJS*, genannt. Mit diesem Verfahren ist es in erster Linie möglich, die *Anzahl* von Netzwerkteilnehmern zu schätzen, die eine gemeinsame Eigenschaft haben. Diese Eigenschaft kann z.B. sein:

- die Existenz (also, dass sie anwesend sind),
- einen Sendewunsch zu haben,
- in einem bestimmten Zustand zu sein.

Das Verfahren arbeitet zunächst nur für eine Gruppe von Teilnehmern in einer Funkzelle, d.h. alle Teilnehmer können mit allen anderen Teilnehmern Daten austauschen ohne weitere Unterstützung. Sie sind paarweise in Funkreichweite.

7.1 Synchrone Jam Signale - SDJS

SDJS ist ein Übertragungsverfahren, das auf überlagerten Funksignalen basiert. Es ist ausschließlich für das Schätzen von Anzahlen entworfen worden und ist nicht als klassische Datenübertragung geeignet. Es werden bei diesem Verfahren keine als Symbole und Signal kodierten Daten ausgetauscht, sondern lediglich Jamsignale versendet. Es findet keine Quellen- oder Datenkodierung, Modulation, etc. statt. Der Geschwindigkeitsvorteil entsteht durch das probabilistische Modell und der Reduktion der Information auf Jamsignale. Funktional (und auch zeitlich) ist das SDJS-Verfahren in drei Schritte unterteilt. In der Klammer ist jeweils angegeben, wer an dem jeweiligen Schritt Anteil hat:

1. Ein Startpaket, das Informationen über den danach folgenden SDJS-Prozess beinhaltet (ein Teilnehmer).
2. Der SDJS-Prozess: Senden und Empfangen von Jamsignalen (alle Teilnehmer).
3. Die Auswertung: Empfangene Jamsignale werden zu Parameterschätzungen verwendet (alle Teilnehmer).

Die technischen Anforderungen an die Hardware und möglichen Implementierungen sind analog zu denen, die in Kapitel 5.4 diskutiert wurden. Im Speziellen sei hier auf Kapitel 5.4.1 verwiesen. Abbildung 7.2 zeigt das SDJS-Verfahren im zeitlichen Ablauf. Das SDJS-Verfahren beginnt mit einem Rundrufpaket, welches die Teilnehmer in der Zelle darüber informiert, wie die Details des folgenden SDJS-Prozesses festgelegt sind. Es enthält außerdem ein Signal für die Zeitsynchronisation. Die Einzelschritte werden nun detaillierter erklärt:

t_1 : Im normalen Datenverkehr wird ein SDJS-Startpaket als Rundruf versendet. Station B ist die Quelle und legt mit diesem Paket die *Parameter* des darauffolgenden SDJS-Prozesses fest. Diese Parameter beinhalten

- die zeitliche Länge der SDJS-Slots,
- die Anzahl der SDJS-Slots,
- die Belegung der Slots.

Ausserdem wird mit Hilfe des Startpaketes auch die Kanalreservierung bis zur Zeit t_7 informiert.

t_2 : Der SDJS-Prozess beginnt. Alle Teilnehmer bereiten ihre Übertragung (den Sendevektor) vor. Dieser wird zufällig gewählt und ist in diesem Beispiel: A : 1001000, B : 1000010 und C : 0010100.

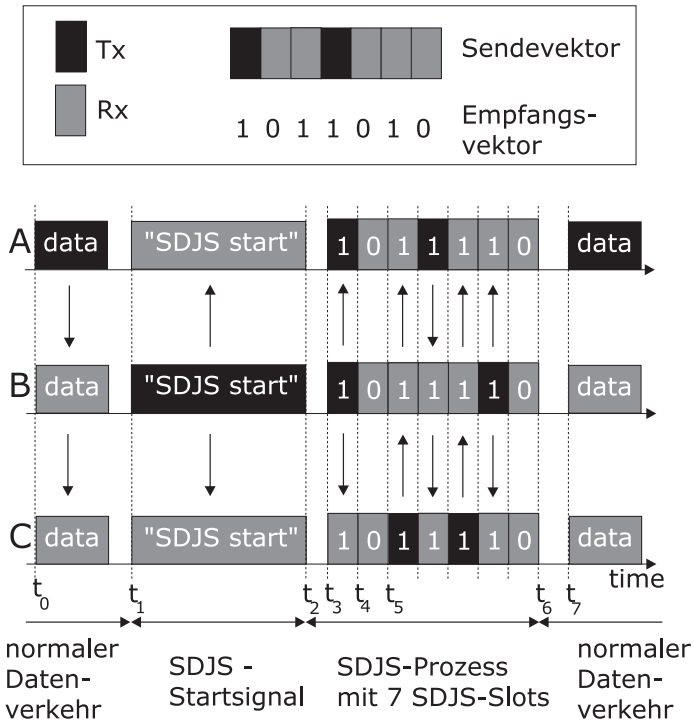


Abbildung 7.2. Die Slotsstruktur des SDJS Verfahrens

- t_3 : Der erste SDJS-Slot beginnt. Teilnehmer A und B wechseln in den Sendemodus und übertragen ein Jamsignal, weil sie eine 1 an der ersten Stelle (im ersten Bit) ihres Sendevektors haben. Teilnehmer C hat eine 0 im ersten Bit seines Sendevektors und schaltet deshalb auf Empfang und sucht Jam-signale. Das Jamsignal (als Summe der Signale von A und B) kann erkannt werden. Teilnehmer C notiert deshalb eine 1 auf der ersten Stelle seines Empfangsvektors. Auch Teilnehmer A und B notieren eine 1 auf der ersten Stelle ihres Empfangsvektors, weil sie selbst gesendet haben und so virtuell ihre eigenen Signale empfangen haben.¹
- t_4 : Alle Teilnehmer haben eine 0 auf der zweiten Stelle ihres Sendevektors. Also schalten sie auf Empfang und versuchen Jamsignale zu entdecken. Da kein Teilnehmer eines sendet, kann nichts erkannt werden. Alle Teilnehmer notieren eine 0 auf der zweiten Position des Empfangsvektors.
- t_5 : Nun hat Teilnehmer C eine 1 auf der dritten Stelle und versendet deshalb ein Jamsignal. Teilnehmer A und B schalten auf Empfang (sie haben eine 0 auf dritter Position des Sendevektors) und sehen das Jamsignal von Teilnehmer C. Daraufhin wird eine 1 auf der dritten Stelle notiert. Dieser Prozess wird genauso bis zur Zeit t_6 fortgeführt.
- t_6 : Der SDJS-Prozess ist beendet und es folgt normaler Datenverkehr. Danach werden die Empfangsvektoren weiter verarbeitet.
- t_7 : Fortsetzung des normalen Datenverkehrs.

Die Verwendung des Empfangsvektors ist sehr vielfältig. Auch die Zusammensetzung dieses Vektors kann sehr unterschiedlich sein. Die oben genannten Parameter des SDJS-Prozesses, die im Startvektor kommuniziert werden, entscheiden auch über die spätere Verwendbarkeit nach dem SDJS-Prozess. In dem diskutierten Beispiel waren die SDJS-Vektoren 7 Slots lang und jede Station hat 2 zufällige Slots belegt. Es gibt natürlich eine Unzahl anderer Wahlmöglichkeiten dieser Bedingungen.

7.2 Parameterschätzung mit SDJS: Anzahl der Geräte

In diesem Abschnitt wird der SDJS-Prozess nun tatsächlich für eine Parameterschätzung eingesetzt. Hierfür werden die notwendigen Modelle und Schätzer her-

¹Der Sendevektor bestimmt ganz allgemein die Abfolge von Senden und Empfangen. In diesem Beispiel ist der Sendevektor für Knoten A 1001000. Mit tx für Senden und rx für Empfangen wird durch den Sendevektor für Knoten A folgender Ablauf festgelegt: $tx-rx-rx-tx-rx-rx-rx$

geleitet [45]. Der Parameter, um den es sich im ersten Beispiel handelt, ist die *Anzahl der aktiven Geräte*. Es soll aus einer Gruppe von Teilnehmern festgestellt werden, wie viele tatsächlich anwesend und aktiv sind. Diese Information ist eine interessante Größe für die Netzwerkschichten, die sich um Verbindung und Transport kümmern. Darüber hinaus ist das hierfür nötige Modell ein Basismodell für den Einsatz und die Implementierung von ähnlichen Schätzaufgaben. Für verschiedene Anwendungen ist die Länge des SDJS-Prozesses variabel, wird jedoch zu Beginn eines SDJS-Prozesses im Startsignal festgelegt.

7.2.1 Analytisches Modell: Das Jäger-Enten-Problem

Für den SDJS-Prozess, bei dem jeder Teilnehmer zufällig (gleichverteilt) einen Zeitschlitz mit einem Jamsignal belegt, können zunächst einige prinzipielle Aussagen getroffen werden:

- Jeder Empfangsvektor wird mindestens eine Eins tragen (Detektionsfehler werden vernachlässigt).
- Jeder Empfangsvektor hat höchstens so viele Einsen wie es Teilnehmer gibt.
- Zwei gleichzeitig sendende Teilnehmer können nicht mehr auseinander gehalten werden.
- Es kann prinzipiell nicht festgestellt werden, ob einer, zwei oder viele in einem Zeitschlitz ein Jamsignal gesendet haben.
- Wenn es deutlich mehr Teilnehmer als Zeitschlitze gibt, so sind wahrscheinlich alle Zeitschlitze mit Jamsignalen belegt, und es kann keine sinnvolle Schätzung mehr durchgeführt werden.

Für die Modellierung werden nun die Bezeichner festgelegt:

- k : die tatsächliche Anzahl der Geräte
- \tilde{k} : die geschätzte Anzahl der Geräte
- A_k : Wahrscheinlichkeitsvariable für die Anzahl der empfangenen Jamsignale (Anzahl der Einsen im Empfangsvektor)
- α : die Anzahl der empfangenen Einsen in einem Experiment
- s : die Anzahl von Zeitschlitzen im SDJS-Prozess (entspricht der Länge der Sende- und Empfangsvektoren)

Die Frage, die das Modell beantworten muss, ist nun: *Wenn bei dem oben festgelegten SDJS-Prozess der Länge s (jeder Teilnehmer wählt nur einen Zeitschlitz*

für ein Jamsignal), die Anzahl der empfangenen Jamsignale a ist, wie viele Geräte k sind dann anwesend? Diese Fragestellung führt zu einem mathematischen Exkurs in ein aus der Kombinatorik bekanntes Analogon:

Das Jäger-Enten-Problem. Eine Gruppe von k Jägern wartet auf einen Schwarm Enten. Sie sind alle erfahrene Jäger; d.h. wenn sie ein Ziel anvisiert haben, treffen sie immer. Jeder von ihnen hat nur eine Patrone Munition. Plötzlich steigt vor ihnen der Entenschwarm in die Höhe. Es sind s Enten. Es geht so schnell, dass die Jäger keine Zeit zur Abstimmung haben und zielen alle zufällig auf eine Ente. Eine gewisse Anzahl a wird getroffen. Da sich die Jäger aber nicht absprechen konnten, werden einige Enten von mehr als einem Jäger getroffen. Die Frage ist nun, wie viele Enten getroffen wurden. Oder genauer: Was ist die Wahrscheinlichkeit, dass a Enten getroffen werden?

Das Modell benutzt das Kombinatorik-Prinzip der Inklusion-Exklusion. Es ist an die Herleitung [76, Theorem 2.1.7] angelehnt.

1. Die Anzahl aller Kombinationen k Jamsignale auf s Slots zu verteilen ist s^k
2. Die Anzahl aller Kombination a Slots aus s Slots zu wählen ist $\binom{s}{a}$
3. Gegeben sind k Elemente. Die Anzahl von Kombinationen, a Gruppen mit jeweils mindestens einem Teilnehmer zu bilden, ist die surjektive Abbildung von $k \rightarrow a$:

$$\sum_{i=0}^a (-1)^i \binom{a}{i} (a-i)^k$$

4. Damit ist die Wahrscheinlichkeit eine Kombination der Größe a zu finden:

$$P(A_k = a|k) = \frac{\binom{s}{a} \sum_{i=0}^a (-1)^i \binom{a}{i} (a-i)^k}{s^k} \quad (7.1)$$

Die Formel (7.1) gibt also die Wahrscheinlichkeit an, dass a Enten sterben, oder analog im SDJS-Prozess a Jamsignale empfangen werden. Abbildung 7.3 zeigt diese Verteilung mit verschiedenen Parametern k . Betrachtet man z.B. die Wahrscheinlichkeitsverteilung für $k = 40$ Geräte, so sieht man, dass der wahrscheinlichste Wert bei $a = 30$ liegt und nicht etwa bei 40. Wenn also 40 Geräte zufällig einen Slot aus 64 belegen, so werden am wahrscheinlichsten nur $a = 30$ Slots belegt sein. Deshalb ist es sehr wahrscheinlich, dass Kollisionen auftreten.

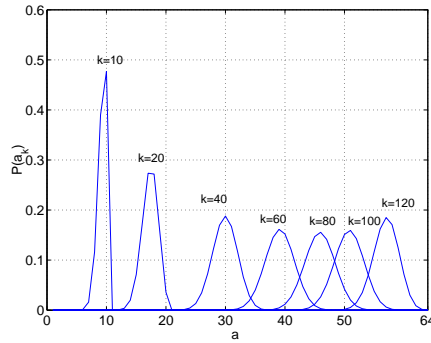


Abbildung 7.3. Wahrscheinlichkeitsverteilung für Jamsignale; $s=64$

Mit Hilfe der Formel (7.1) kann man verschiedene Schätzer herleiten. Es bieten sich der *Maximum-Likelihood (ML)*-Schätzer:

$$\tilde{k}_{ML} = \arg \max_k P_{a|k}(a|k) \quad (7.2)$$

oder der *Maximum a posteriori (MAP)*-Schätzer an:

$$\begin{aligned} \tilde{k}_{MAP} &= \arg \max_k P_{k|a}(k|a) \\ &= \arg \max_k \frac{P_{a|k}(a|k) \cdot P_k(k)}{P_a(a)} \\ &= \arg \max_k P_{a|k}(a|k) \cdot P_k(k) \end{aligned} \quad (7.3)$$

Um den MAP-Schätzer einzusetzen, ist es notwendig, die a-priori-Verteilung $P_k(k)$ zu kennen. In den meisten Fällen ist es jedoch schwer oder unmöglich, hier eine Verteilung anzugeben. Es kann jedoch bereits hilfreich sein, Charakteristika der Verteilung zu erfassen und unbekannte Bereiche einfach gleichverteilt anzunehmen. Nimmt man die gesamte Verteilung $P_k(k)$ als gleichverteilt an, so erhält man wieder den ML-Schätzer.

7.2.2 Maximum-Likelihood-Schätzer

Der Maximum-Likelihood-Schätzer basierend auf (7.2) ist in Abbildung 7.4 beispielhaft dargestellt. Die Parameter waren $s = 32$ und $k \in [0, 120]$. Die Abbil-

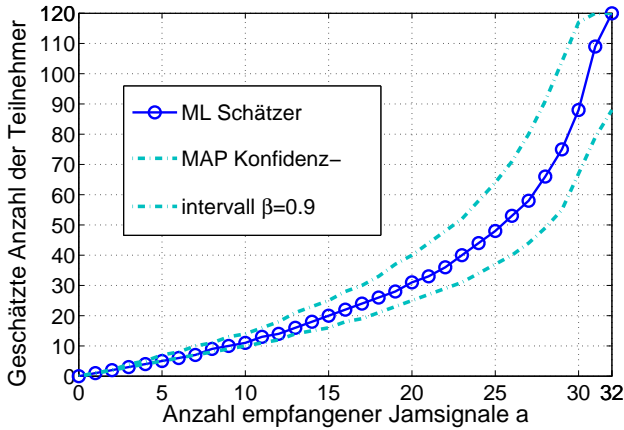


Abbildung 7.4. ML- und MAP Schätzer für $s = 32$

Abbildung zeigt den ML-Punktschätzer, der jedem Wert für a einen Wert für die Schätzung zuordnet. Die Schätzung kann also als Tabelle abgelegt werden. Die MAP-Konfidenzintervalle (hier mit $P_k(k) = \text{const}$) geben denjenigen Bereich an, aus dem 90% ($\beta = 0.9$) der möglichen Geräteanzahlen sind, die nach einem SDJS-Prozess auf das jeweilige a abbilden. D.h. für e.g. $a = 20$ ist aus $k \in [0, 120]$ die Wahrscheinlichkeit 0.9, dass $\tilde{k} \in [25, 40]$. Nimmt man für $P_k(k)$ kein weiteres Wissen an, so würde der ML-Schätzer für $a = s$ stets $\tilde{k} = \infty$ schätzen. Deshalb sollte man hier stets eine sinnvolle Obergrenze annehmen, um so auch für $a = s$ sinnvolle Schätzwerte zu erhalten. In Abbildung 7.5 ist ein ML-Schätzer mit solchen Obergrenzen abgebildet. Es ist genauer gesagt ein MAP-Schätzer mit einer vorgegebenen Gleichverteilung auf den Intervallen $k \in [k_{\min} = 0..k_{\max} = 50, 70, 120]$. Interessant ist hier, dass für die unteren Bereiche von a kein Unterschied zwischen MAP-Schätzern besteht, wenn die Originalverteilung eine Gleichverteilung ist. Der Schätzer ist also invariant, solange der geschätzte Wert innerhalb des erlaubten Bereiches $[k_{\min}..k_{\max}]$ liegt.

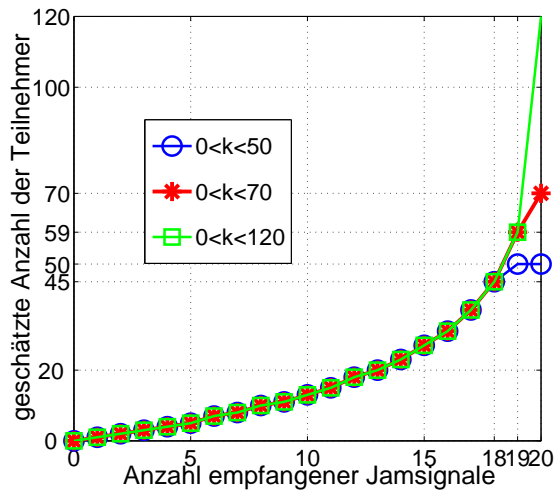


Abbildung 7.5. Invarianz des MAP/ML-Schätzers

7.2.3 Implementierung und Evaluierung

In der Simulationsimplementierung wurde ein SDJS-Prozess mit 256 Slots evaluiert. Es wurde ein ML-Schätzer basierend auf einer Schätztablelle verwendet. Darüber hinaus wurden symmetrische Kanalfehler mit Fehlerraten der Entdeckung und Fehlalarme von bis zu $\text{SlotErrorRate}(\text{SER}) = 10^{-2}$ eingesetzt. Eine solche Annahme ist weitaus pessimistischer als normale Fehlerquoten, die für drahtlose Netzwerksimulationen angenommen werden². Insgesamt über 20000 simulierte Schätzungen bilden die Basis der Graphik 7.6. Die y-Achse gibt die akkumulier-

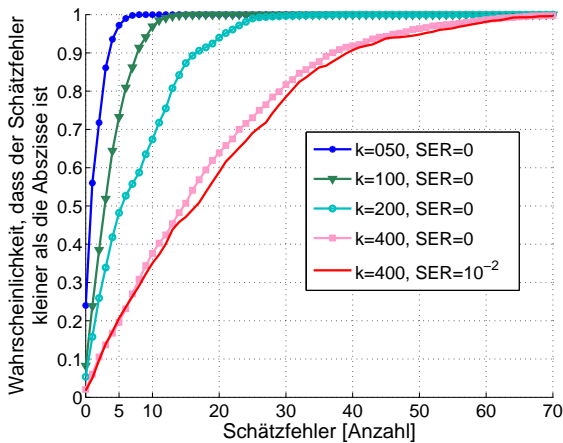


Abbildung 7.6. SDJS-Simulation mit $s = 256$

ten Wahrscheinlichkeiten für das Auftreten von Schätzfehlern an. Für $k = 100$ haben z.B. 95% der Simulationen einen Fehler kleiner oder gleich 8 ergeben. Für $k = 400$ ist auch noch eine Kurve mit $\text{SER} = 10^{-2}$ angegeben. Man sieht, dass SDJS diese Kanalfehler sehr gut verkraften kann. Insgesamt ist jedoch auch ein deutlicher Trend festzustellen. Für größere k ist die Schätzung insgesamt schlechter. Das liegt in der Natur der SDJS-Schätzverfahren: Die Schätzung ist für hohe Zahlen schlechter als für kleine, da es zu mehr Kollisionen kommt. Dieser

²typischerweise 10^{-3}

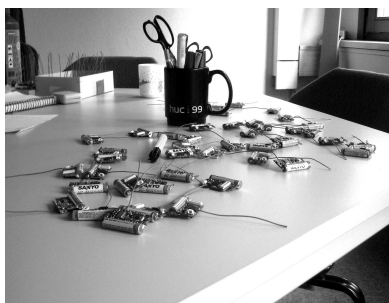


Abbildung 7.7. Ein Experiment mit über 40 Particle Computern

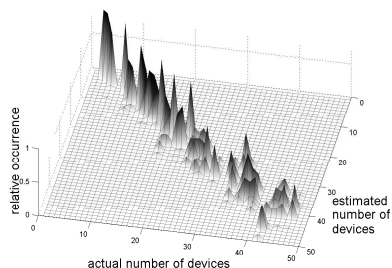


Abbildung 7.8. Darstellung der Schätzergebnisse des Experiments mit 40 Particle Computern

nichtlineare Zusammenhang ist auch in Abbildung 7.4 daran zu sehen, dass die Konfidenzintervalle für ansteigende Werte von a übermäßig größer werden.

7.2.4 Experiment und Verifikation

Das SDJS-Verfahren wurde auf der Hardware Plattform Particle Computer [13] [17] implementiert. Die Länge eines SDJS-Slots war bei der Implementierung $64\mu s$. AwareCon benutzt ein TDMA-System mit einem festen Zeitrahmen von 13ms. Der SDJS-Prozess mit 64 Slots wurde in diesem Zeitrahmen eingebettet. Das Testszenario, wie es auf Abbildung 7.7 zu finden ist, bestand aus 5 bis zu 50 Particle Computern, die alle in Funkreichweite in einem Büroraum verteilt wurden. Insgesamt wurden über 1820 Einzelmessungen durchgeführt. Das Ergebnis dieser Experimente ist in Abbildung 7.8 zu sehen. Die z-Achse zeigt die Häufigkeit des Auftretens eines x-y-Paares „Tatsächliche Anzahl - Geschätzte Anzahl“. Somit bekommt man einen guten visuellen Eindruck der Qualität des Verfahrens. Die Schätzung folgt – wie erwartet – der Winkelhalbierenden von x- und y-Achse. Für größere Anzahlen wird die Verteilung „breiter“. Die Fehler wurden in einer weiteren Graphik (Abbildung 7.9) zusammengefasst. Hier sieht man, dass z.B. für $k = 20$, der Fehler in 95% der Fälle unterhalb von 3 Geräten lag. Für $k = 30$ enthält diese Graphik auch noch eine Kurve, die aus der Simulation mit $SER = 10^{-2}$ gewonnen wurde. Es ist der beeindruckende Beweis, dass die Implementierung und

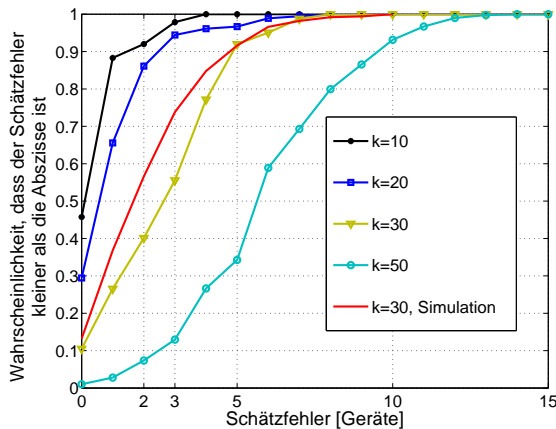


Abbildung 7.9. Akkumulierte Fehlerverteilung im Experiment mit 40 Particle Computern; Vergleich mit der Simulation

die Simulation die gleichen Ergebnisse liefern und dass SDJS der Überprüfung in einer echten Hardwareimplementierung standgehalten hat.

7.2.5 Vorteile der Parameterschätzung mit SDJS

Das Protokoll SDJS wurde als Lösungsvorschlag des Problems der Schätzung von Parametern entworfen, wenn die nötige Information verteilt vorliegt. Hier kann SDJS gegenüber herkömmlichen Verfahren deutlich schneller und effizienter arbeiten. Die Vorteile von SDJS für die Parameterschätzung sind im Einzelnen:

Einfache Implementierung. Der Einsatz von SDJS beruht auf dem Senden und Detektieren von Jamsignalen. Diese können mit einfachen Mechanismen wie Senden des unmodulierten Trägers und Trägererkennung bereits von Standardhardware geleistet werden.

Schätzung über eine Tabelle. Die Schätzung der Parameter mit SDJS kann in vielen Fällen über eine Schätztabelle realisiert werden. So ist keine aufwendige Softwareimplementierung auf den Sensorknoten nötig.

Geschwindigkeit. SDJS ist durch seine parallele Verarbeitung deutlich schneller als bekannte Mechanismen. Bei der Schätzung der Anzahl der Geräte kann im Vergleich zu klassischen Mechanismen der Prozess in einem Bruchteil der Zeit durchgeführt werden. Der Vorteil kann leicht den Faktor 100 überschreiten.³

7.3 Parameterschätzung mit SDJS: Anwendung im Supermarkt

Der Supermarkt ist ein interessanter Bereich für den Einsatz von Monitoring oder Identifikationstechnologie. Die Prozesse und Arbeiten umfassen viele Domänen — von einfachen manuellen Tätigkeiten bis hin zu Informationsverarbeitung für Marketing und Rechnungswesen. Es gibt bereits einige technische Lösungen, die auch die Interaktion mit den Kunden unterstützen. Ein Beispiel sind sogenannte *Recommender-Systeme*. Solche Systeme informieren den Kunden über Angebote und Sortiment über variable Displays oder Sprachausgabe. Sie interagieren mit dem Kunden über elektronische Medien oder direkt während des Einkaufs im Supermarkt. Die meisten technischen Systeme zielen jedoch darauf ab, die Personalkosten im Supermarkt zu reduzieren. Das bekannteste Beispiel ist hier die automatische Kasse. Erste Initiativen haben z.B. Metros *Extra Future Store* [3], [56] in Rheinberg hervorgebracht, wo RFID-Technologie bereits für automatische Kassen im Einsatz ist. Um einen besseren Einblick in die Prozesse zu bekommen, wurde eine Studie mit *dm-drogerie markt*⁴ in Karlsruhe und Umgebung durchgeführt. Basierend auf dieser Studie wurde ein Prozess im Supermarkt identifiziert, der geeignet war, um technologische Fortschritte mit überlagerten Signalen zu diskutieren [46]. Es handelt sich dabei um das Überprüfen von Mindesthaltbarkeitsdaten. Dieser Prozess ist immer noch sehr schwierig und auch mit RFID-Technologie nicht ausreichend effizient.

7.3.1 Einsatz von Informationstechnologie im Supermarkt

Heutzutage ist Informationstechnologie im Supermarkt hauptsächlich darauf ausgerichtet, Produkte zu *klassifizieren*. Das Referenzbeispiel sind die Standards für

³Beispielrechnung für pParts: pro 13ms-Zeitrahmen kann ein Paket verschickt werden. Bei 200 Knoten braucht man zum Sammeln der IDs im Idealfall $200 * 13ms = 2600ms$. SDJS kann auf pParts mit einer SDJS-Slotzeit von $64\mu s$ arbeiten. In einem 13ms Zeitrahmen wurden 64 SDJS Slots realisiert. Somit ist der Zeitvorteil hier $2600 : 13 = \text{Faktor } 200$.

⁴<http://www.dm-drogeriemarkt.de>

Barcodes, wie sie im EAN (European article numbering system) oder UPC (universal product code) festgelegt wurden. Eine solche Klassifizierung ist ausreichend, um Prozesse wie Nachbestellungen oder Kassieren zu unterstützen.

Der nächste technologische Schritt nach dem Barcode ist RFID. Mit RFID können Produkte über Funk erfasst und identifiziert werden. Damit ist ein entscheidender Wechsel von *Klassifikation* hin zu individueller *Identifikation* vollzogen. Auch wenn die meisten Anwendungen mit RFID noch in der Kinderschuhen stecken und die Technologie auch noch nicht kostengünstig genug ist, stimmen alle Analysten darüber ein, dass RFID in naher Zukunft besonders für den Einzelhandel ein enormes Potenzial hat.

Ein Problem, das mit der individuellen Identifikation einhergeht, ist die Datenexplosion. Die eindeutige Identifikation eines Warenstücks benötigt eine virtuelle Repräsentation in den elektronischen Datenverarbeitungssystemen. Um diese Datenflut zu vermeiden ist es absehbar, dass Produkte in Gruppen oder Klassen zusammengefasst werden. Die hohe Anzahl der Produkte führt auch zu Schwierigkeiten in den Funkprotokollen bei z.B. dem Kanalzugriff. Hier sollen überlagerte Funksignale eine Lösung bieten, die über Standardansätze von RFID-Protokollen hinausgeht. Basis für die Architektur ist das SDJS-Verfahren aus Abschnitt 7.2.

7.3.2 Studien im Supermarkt

Um den Einsatz von neuartiger Informationstechnologie im Supermarkt zu beurteilen, wurde eine Studie mit dm-drogerie markt durchgeführt. In dieser Studie sollte vor allem der Einsatz von RFID oder ähnlicher Technologie für besondere Aufgaben im Supermarkt untersucht werden. Es wurden besondere manuelle Tätigkeiten analysiert und eine Befragung von vier Ladenmanagern und einigen Angestellten durchgeführt. Die Studie wurde außerdem von sechs Managern aus den Hauptniederlassungen unterstützt. Diese Studie [46] umfasst Fragen zur Thematik des manuellen Prozesses für die Überprüfung der Mindesthaltbarkeitsdaten von verderblicher Ware. Fragen waren z.B.

- Wie oft werden die Regale nach abgelaufener Ware überprüft?
- Nach welchen Kriterien werden Produkte vom Regal genommen?
- Die lange dauert es, eine Produktgruppe zu überprüfen?

Diese Fragen führten zu einer Klassifizierung in drei Produktklassen. Die erste Klasse besteht aus Produkten ohne Mindesthaltbarkeitsdatum; diese benötigen

keine Überprüfung. Die zweite Klasse besteht aus Produkten mit Mindesthaltbarkeitsdatum; diese benötigen regelmäßige Überprüfung. Die dritte Klasse sind Produkte, die teilweise Mindesthaltbarkeitsdaten tragen. Die Zeit für die Überprüfung hängt sehr von der Vertrautheit des Angestellten mit den Produkten ab. Darüber hinaus ist die nötige Zeit für die Überprüfung auch sehr von der Form und Beschaffenheit der Produkte abhängig. Nach den Zahlen der Studie liegt die Überprüfungszeit eines einzelnen Produktes zwischen *sieben* und *vierzehn* Sekunden. Die Quote der Produkte, die aussortiert werden müssen, liegt typischerweise bei 0,25%. Es liegt also nahe, eine technische Lösung anzustreben, die den enormen Aufwand reduziert, der benötigt wird, um diese 0,25% aufzudecken. In Abbildung 7.10 sind die Ergebnisse der Befragung zu den manuellen Prozessen zusammengefasst. Aus den Zahlen der Befragung wurden die mittleren Überprüfungszeiten für die Produktgruppen bestimmt.

	different products on the shelf	items on the shelf	time to check the shelf	check time per item
whole food products	241	3000	6h	7sec
pharmacological products	487	4000	16h	14sec

Abbildung 7.10. Einige Zahlen über den manuellen Überprüfungsvorgang. Aus [46, Abbildung 2]

Die Überprüfungen werden normalerweise regelmäßig nach einem Arbeitsplan erledigt. Es kann kein individueller Prozess für einzelne Produkte erreicht werden, obwohl Umwälzung und Haltbarkeit der Produkte äußerst verschieden sind. Es wäre also sehr interessant, eine Überprüfung nur bei Bedarf zu durchzuführen, denn oft wird kein einziges abgelaufenes Produkt gefunden. Darüber hinaus sind die Manager auch daran interessiert, den Verlauf der Mindesthaltbarkeitsdaten (MHD) einzelner Produktgruppen zu beobachten, um mit Marketing und Sortimentsoptimierungen das Angebot zu verbessern.

7.3.3 Multi-SDJS für MHD-Profile auf RFID

Die Idee ist nun, ein System zu entwerfen, welches diesen Prozess der Überprüfung der Mindesthaltbarkeitsdaten vereinfacht. Der Vorschlag ist die Nutzung von

Profilen. Diese Profile geben einen Überblick über die Verteilung von Mindesthaltbarkeitsdaten einer bestimmten Gruppe von Produkten. Abbildung 7.11 zeigt

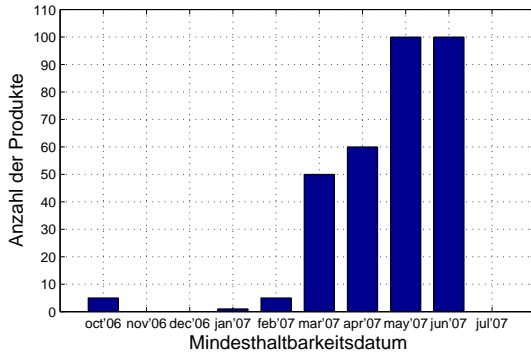


Abbildung 7.11. Profil von Mindesthaltbarkeitsdaten

ein solches Profil von Mindesthaltbarkeitsdaten. Die Auflösung in der Zeit ist eine Frage der Implementierung und ist beliebig wählbar. Die Technologielösung mit RFID und SDJS kann diese Profile durch kollaborative Signalisierung generieren. Mit Hilfe von SDJS ist der Zeitaufwand, ein Profil einer Gruppe zu erzeugen, gegenüber bekannten RFID-Protokollen deutlich beschleunigt, so dass er in einer Zeit von weniger als 50ms durchführbar ist.

Voraussetzung für den Einsatz von SDJS für dieses Szenario ist, dass alle Produkte mit RFID-Label versehen sind. Dies ist zwar heutzutage noch nicht der Fall, es macht dennoch Sinn, die heutigen Protokolle schon jetzt einer kritischen Prüfung zu unterziehen, damit durch alternative Lösungen nicht zuletzt auch RFID schneller in Endanwendungen eingesetzt werden kann.

Lösungen basierend auf RFID-Standards

Wenn man die oben diskutierten Profile mit Hilfe von Standard-RFID-Protokollen implementieren wollte, so ist dies prinzipiell möglich.

Harald Vogt hat in einer Veröffentlichung [81] den Identifizierungsprozess von einer Gruppe von RFID-Etiketten (engl. „tags“) genauer untersucht. Legt man nun die Ergebnisse dieser Veröffentlichung zu Grunde, so würde ein Identifikationsprozess von einer Gruppe von 80 tags über 8000ms liegen, um eine 99%-ige Erkennung aller Tags zu erzeugen. Für die Zeitberechnung wird hier die serielle Anbindung des Lesegerätes mit 56kBit/s verwendet. Weitere Betrachtungen zeigen dort, dass eine Gruppe von 200 Tags mehr als $256 \cdot 13$ Slots benötigt, um eine 99%-ige Erkennung zu bekommen. Hierbei entspricht ein Slot einem Datenpaket. Wenn man nun für die Zeitberechnung eine 64-bit Identifikation und die PHY-Parameter des schnellsten RFID-Standards (ISO18000-3 Mode 2 [38], Slotzeit ist $14.2\mu s$) zu Grunde legt, dann erreicht man eine Gesamtzeit von 3000ms ohne Overhead durch Nachrichten-Header und Kontrollpakete vom Lesegerät. Dies wäre also das theoretische Minimum unter dem Modell von Vogt und den Parametern der physikalischen Schicht von ISO18000-3 Mode 2.

Andere Verbesserungsvorschläge zu RFID-Protokollen wurden zum *Mode 5* beigetragen; dieser ist aber wegen der hohen Hardwareanforderungen nicht zum Standard erhoben worden. In [65], [2] behaupten die Hersteller von RFID-Hardware, dass sie 1200 Tags pro Sekunde identifizieren können. Dies bezieht sich allerdings auf ganz spezielle Szenarien. Hier dreht es sich stets um mobile Szenarien, in denen während einer bestimmten Zeit eine bestimmte Anzahl von Tags nacheinander durch ein Tunnellesegerät gefahren wird. In statischen Szenarien sind die Lesegeschwindigkeiten typischerweise als 100-200 Labels pro Sekunde bekannt. Darüber hinaus sind auch der Identifikationsprozess und das Datenauslesen zwei unterschiedliche Prozesse. Um ein Mindesthaltbarkeitsdatum zu übertragen, müsste man also nach der Identifikation auch noch die Daten aller Tags auslesen.

Um die Schwächen von existierenden RFID-Protokollen auszugleichen, soll nun basierend auf SDJS ein Protokoll vorgeschlagen werden, mit dem die Mindesthaltbarkeitsdaten von Produkten schnell und effizient ausgelesen werden können. Ziel ist es dabei nicht, ein allgemeines Protokoll für Datenübertragung zu entwerfen. Es ist eine spezielle Lösung, die nur für ganz bestimmte Datenprozesse geeignet ist.

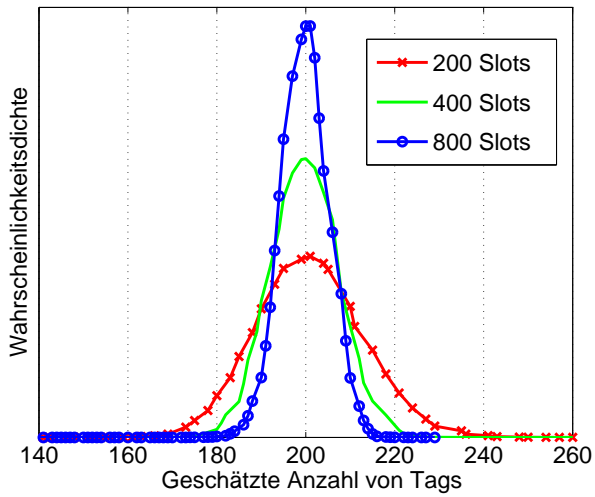


Abbildung 7.13. Wahrscheinlichkeitsdichte für $k = 200, s = 200, 400, 800$

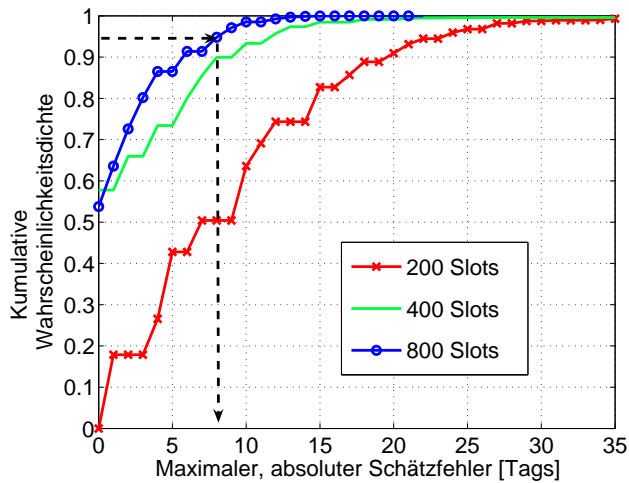


Abbildung 7.14. Akkumulierte Wahrscheinlichkeitsverteilung für den Schätzfehler bei $k=200$ Tags

Parameter (also der Anzahl der Slots) kann man die Genauigkeit von SDJS festlegen. Genauer sieht man dies noch einmal in Abbildung 7.14. Da die Schätzungen stets auf einer Tabelle mit diskreten Werten basiert, sind diese diskreten Stufen auch in der Genauigkeitsanalyse wiederzufinden. In Abbildung 7.13 wurden diese zur besseren Lesbarkeit graphisch geglättet. Wenn man nun mit Hilfe von Abbildung 7.14 aussagen will, welche Genauigkeit bei einem gegebenen Prozess zu erwarten ist, so muss man zunächst eine Konfidenz der Aussage wählen. In diesem Beispiel wurde eine Konfidenz von 95% gewählt. Der Pfeil in Abbildung 7.14 startet also bei Ordinate 0,95. Für das Szenario mit 800 Slots erreicht man dann einen Fehler von höchstens 7 Tags. Zusammengefasst heißt das, dass bei SDJS Prozessen mit 200 Tags und 800 Slots in 95% der Fälle der Schätzfehler höchstens 7 Tags ist.

7.3.4 Implementierung

Die Implementierung wurde wieder auf pParts Particle Computer durchgeführt. Hierzu wurde das proprietäre Protokoll AwareCon so angepasst, dass es das Verhalten von RFID-Protokollen beim Kanalzugriff nachahmt. Das Lesegerät wurde ebenfalls durch einen Particle Computer ersetzt, der die entsprechenden Befehle als Rundruf an die Tags versendet hat. Als Beispiel wurde eine beliebige Verteilung von Mindesthaltbarkeitsdaten in die Tags einprogrammiert. Nach dem Startsignal des Lesegerätes antworteten die Tags mit dem oben diskutierten Multi-SDJS Verfahren. Mit Hilfe einer Schätztabelle wurde dann die Anzahl von Produkten mit den jeweiligen Mindesthaltbarkeitsdaten angegeben. Im Experiment wurden 50 pParts in einem Büro in ein Regal verteilt, und der Auslesevorgang wurde angestoßen. Das Übertragungsverhalten von RFID auf der Funkschnittstelle kann mit den pParts nicht vollständig dargestellt werden. Speziell die Übertragung im Kanal ist bei passiven RFID-Tags deutlich unterschiedlich zu pParts. Dies ist jedoch für das vorgeschlagene Multi-SDJS Protokoll nicht kritisch. Im Lesegerät müssen lediglich Jamsignale erkannt werden. Dies entspricht für die RFID-Tags dem Senden eines Signals — oder genauer gesagt: dem Rückkoppeln von eingestrahelter Energie bzw. dem Verzerren des Feldes durch Resonanzkurzschluss.

In Abbildung 7.15 ist das Ergebnis eines Experiments zu sehen. Die geschätzten Zahlen von SDJS stimmen sehr gut mit den tatsächlichen überein. Nur an der Position „jun“ ist ein Schätzfehler von zwei Tags entstanden.

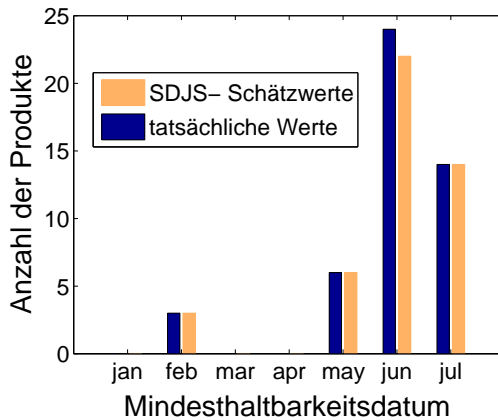


Abbildung 7.15. Profil der Mindesthaltbarkeitsdaten für 50 Produkte beim Experiment mit der Implementierung

7.3.5 Zusammenfassung

Das in Abschnitt 7.3.3 vorgeschlagene Verfahren kann die Erstellung der Profile von Mindesthaltbarkeitsdaten deutlich beschleunigen. In einem Rechenbeispiel sollen nun die Ergebnisse von Vogt [81] und die Ergebnisse von SDJS verglichen werden. Dafür wird ein Szenario von 200 Tags betrachtet. Mit den theoretischen Ergebnissen aus [81] und den physikalischen Parametern aus [38] kann man die optimale Lesezeit hierfür auf 3000ms schätzen. Diese Zeit berücksichtigt jedoch nicht den Overhead von Protokoll-Headern, Kommunikation mit dem Datenverarbeitungssystem etc. Es ist eine theoretische, untere Grenze, die von existierenden Systemen nie erreicht wird. Das SDJS-Verfahren würde mit einer Unterteilung in 48 Monate und 100 Slots für einen Monat und $10\mu s$ pro SDJS-Slot eine Gesamtzeit von 49ms benötigen. Das ist eine 6000%-ige Verbesserung gegenüber Standard-RFID-Protokollen. Darüber hinaus wird bei SDJS die benötigte Zeit nicht mehr, wenn die Anzahl der RFID-Tags steigt. Die Dauer des Prozesses hängt nur von der Anzahl der Unterteilungen, also der gesamten Anzahl der SDJS-Slots ab.

7.4 Multi-SDJS: Datenfusion von Temperaturmessungen

In den vorangegangenen Kapiteln wurden SDJS, Multi-SDJS und Anwendungen, die Anzahlen von Geräten oder Objekten schätzen, vorgestellt. Nun soll noch eine Erweiterung von SDJS zur Schätzung von *Mittelwerten* vorgestellt werden. Es sollen also verteilte Messwerte r_i eingesammelt und deren Mittelwert

$$\hat{r} = \frac{1}{N} \sum_{i=0}^{N-1} r_i \quad (7.4)$$

berechnet werden. Um dies besonders effizient und schnell zu tun, soll wieder SDJS eingesetzt werden. Da SDJS oder auch Multi-SDJS nur Anzahlen schätzen kann, muss der Prozess der Mittelung (7.4) so verändert werden, dass er für SDJS geeignet ist. Dafür wird der Wertebereich von $r \in [r_{min}, r_{max}]$ in M feste Intervalle $I_j = [r_j, r_{j+1}]$ unterteilt. Für jedes dieser Intervalle wird nun die Anzahl \tilde{N}_j der Messwerte, die darin liegen, mit SDJS geschätzt. Dies ist nichts weiter als die Verallgemeinerung des bereits vorgestellten Multi-SDJS-Protokolls. Im Startpaket für dieses Multi-SDJS-Protokoll müssen also die Intervalle kommuniziert werden. Die Mittelung sieht dann wie folgt aus:

$$\hat{r}_{SDJS} = \frac{1}{\sum_{j=0}^{M-1} \tilde{N}_j} \sum_{j=0}^{M-1} \underbrace{\frac{r_j + r_{j+1}}{2}}_{\text{Schwerpunkt des Intervalls}} \cdot \tilde{N}_j \quad (7.5)$$

Die Quantisierung, die durch die Einführung der Intervalle geschieht, erzeugt natürlich in der Mittelung Fehler. Wenn die Intervalle zu groß gewählt werden — im Extremfall z.B. ein Intervall, das sämtliche Messwerte enthält —, so ist das Ergebnis kaum noch von den tatsächlichen Werten beeinflusst. Auch die Anzahl der SDJS-Zeitschlitze für jedes Intervall ist wichtig, da es die Schätzqualität von SDJS beeinflusst. Wichtig ist hier auch zu bemerken, dass mit diesem Verfahren nur *skalare* Größen verarbeitet werden können. Mehrdimensionale Werte müssten also komponentenweise verarbeitet werden, falls dies möglich ist.

Konzeptnachweis

In den Abschnitten 7.2.4 und 7.3.4 wurden schon Implementierungen von SDJS und Multi-SDJS vorgestellt. Deshalb soll die Diskussion über Implementierungen an dieser Stelle kurz gehalten werden. In der Implementierung wurden hier

eine SDJS-Slotlänge von $64\mu s$ und 64 Slots pro Werteintervall gewählt. Es gab insgesamt 14 solcher Intervalle. In der Testimplementierung für die Fusion von Messwerten wurde ein Experiment mit ca. 30 pPart Particle Computern durchgeführt. Jeder der Knoten verhielt sich auf SDJS-Protokollebene so wie zwei Knoten, was insgesamt zu über 50 virtuellen Teilnehmern führte. Als Beispiel wurde hier die Temperatur gemessen, und die Messwerte bewusst mit einem Rauschfehler versehen. Abbildung 7.16 zeigt das Experiment. Die Messwerte sind in Ab-

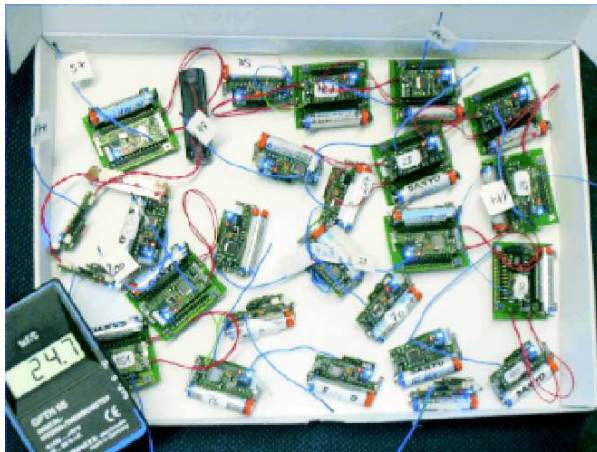


Abbildung 7.16. Einsammeln und Mittelung mit dem Multi-SDJS-Prozess: mehrere Sensorknoten mitteln ihre Temperaturmessung

bildung 7.17 zu sehen. Man kann gut erkennen, dass durch SDJS Schätzfehler für die einzelnen Anzahlen auftreten. Vergleicht man allerdings den wahren Mittelwert mit demjenigen, der von SDJS bestimmt wurde, so sieht man, dass zwischen $\mu_{real} = 25,0$ und $\mu_{SDJS} = 25,3$ kein großer Unterschied besteht. Die wahre Temperatur in diesem Experiment wurde zu $24,7^{\circ}C$ bestimmt, somit ist der Fehler nur 2,3%. Dieses Experiment zeigt, dass das Multi-SDJS-Verfahren auf Intervallen von Messwerten gut geeignet ist, um z.B. den Mittelwert von vielen, verteilten Messungen schneller und effizienter zu bestimmen.

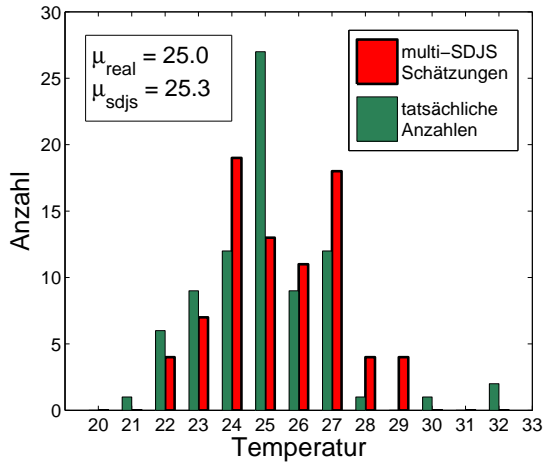


Abbildung 7.17. Der Multi-SDJS-Prozess für das Einsammeln von Daten

7.5 Parameterschätzung mit SDJS: Datenfusion für Lokationssysteme

Die Kenntnis der physikalischen Topologie von Sensorknoten ist für viele Anwendungen wichtig. Diese reichen von Netzaufgaben wie Routing und Sendeleistungsregelungen über Annotation von Sensorwerten bis hin zu Bereichen wie Mobile Computing und Ubiquitous Computing [22]. Viele Lösungen für die Lokalisation von Sensorknoten wurden vorgeschlagen, die sowohl die Messtechnik als auch die Algorithmik behandeln. Viele Arbeiten befassen sich damit, die Genauigkeit der Ortsbestimmung der Sensorknoten zu verbessern, während der Berechnungsaufwand in Schranken gehalten wird. Die Fehler, die auf den rohen Messwerten liegen, beeinflussen dabei die Genauigkeit der Ortsbestimmung erheblich. Dies wurde z.B. in [84, 69] untersucht. In diesem Kapitel soll gezeigt werden, wie die systematischen Messfehler solcher Lokationssysteme deutlich verringert werden können. Der Einsatz von überlagerten Funksignalen (speziell des SDJS-Protokolls) sorgt dafür, dass das gesamte System skalierbar bleibt und

ein Masseneffekt genutzt werden kann: je mehr Knoten an dem System teilnehmen, desto besser sind die Ergebnisse.

7.5.1 Andere Arbeiten im Bereich Lokalisierung von Sensorknoten

Lokationssysteme haben viele Anwendungsbereiche. So gibt es z.B. Systeme zur Verfolgung von Menschen und Objekten [22]. Diese haben eine Vielzahl von Medien zur Messung untersucht. Beispiele sind Infrarotsignale, Ultraschall, Empfangsleistungen von Funksignalen, Bildverarbeitung und physikalischer Kontakt. Die meisten dieser Systeme benötigen jedoch Infrastruktur und erwarten eine statische Installation.

Systeme, die speziell für Sensornetzwerke geeignet sein sollen, müssen die Einschränkungen in der Prozessleistung, Kommunikation, Energiereserven und Mobilität berücksichtigen. Die meisten dieser Systeme basieren auf der Messung von Funkempfangsleistungen (RSSI; radio signal strength indication) [23, 86], oder Laufzeitmessungen von akustischen Signalen [70, 63, 58]. Lokationssysteme sind Gegenstand aktiver, gegenwärtiger Forschung. In jüngster Vergangenheit wurde gezeigt, dass präzise Abstandsmessungen zwischen Knoten möglich sind, indem Phasenverschiebungen zwischen zwei nahe beieinander stehenden Trägerfrequenzen bestimmt werden [54]. Andere Autoren haben vorgeschlagen, den Empfangswinkel (AoA, angle of arrival) zu nutzen, um die Ausrichtung [63] und letztlich auch den Ort von Sensorknoten zu bestimmen [61].

Auf Algorithmenseite gibt es ebenfalls Bemühungen, um die bestmögliche Verarbeitung der Messwerte für eine möglichst genaue Ortsbestimmung der Knoten zu erreichen [49]. Einige Algorithmen benötigen keine Ankerknoten, deren Ort a priori bekannt ist. Solche Algorithmen erzeugen dann *relative* Koordinaten, die die Lage der Knoten untereinander ohne globale Referenz beschreiben [80, 72]. Auch die Bedeutung von Messfehlern und Rauschen auf den Messwerten wurde bereits ausgiebig untersucht [59, 84]. In [69] zeigen die Autoren theoretische Grenzen für die Genauigkeit der Ortsbestimmung, wenn die Größe der Messfehler bekannt ist.

7.5.2 Anwendungen und Motivation

Die weiteren Betrachtungen sollen nun auf Szenarien fokussiert werden, in denen besonders hohe Knotendichten vorkommen, und besonders hohe Anzahlen

von Knoten beteiligt sind. In der Literatur findet man viele Arbeiten, die sich mit solchen Szenarien beschäftigen [49, 72, 69]. Beispiele sind hier die Überwachung von Umweltparametern oder militärische Anwendungen. Ein weiteres Anwendungsgebiet ist das von Ubiquitous Computing [82]. Alltagsgegenstände, die mit Kleinstrechnern ausgestattet werden, erzeugen schnell eine Umgebung, in der hunderte Funkknoten in gegenseitiger Reichweite sind. Diese Gegenstände verarbeiten oft Kontextdaten — einer der wichtigsten Kontexte ist die Lokation. Deshalb ist eine genaue Ortsbestimmung eine wichtige Voraussetzung für viele Anwendungen von Ubiquitous Computing.

Das System, das nun vorgeschlagen wird, nutzt die Eigenschaft genau dieser Umgebungen: die große Anzahl von Sensorknoten soll die Genauigkeit erhöhen und das Gesamtergebnis, nämlich die Ortsbestimmung der einzelnen Sensorknoten, verbessern. Entscheidend ist hierbei die Idee des *Masseneffekts*: Je mehr Knoten an dem System teilnehmen, desto besser soll die Lokalisation werden.

7.5.3 Methodik und Fehlermodelle

Wie in vielen genannten Veröffentlichungen, dreht sich auch dieses System um die Analyse von Messfehlern. Es wird jedoch ein neuer Weg beschritten: anstatt zu versuchen, immer genauere Messungen und Algorithmen zu erzeugen, sollen in diesem System die *systematischen* Messfehler diskutiert werden, die in vielen Lokationssystemen vorkommen. *Systematische* Fehler sind diejenigen Fehler, die für eine feste Konstellation von Hardware und Umgebung wiederholbar sind. Sie können also nicht durch mehrfache Messungen ausgeglichen werden. Für Lokationssysteme sind diese systematischen Fehler besonders vom Ort und von der Orientierung der Objekte abhängig. Deshalb wird vorgeschlagen, Messungen der gleichen physikalischen Größe von *verschiedenen* Knoten zusammenzuführen, um so die *verschiedenen* systematischen Fehler der Knoten in einen Datenfusionsprozess einzuleiten. Der systematische Fehler einer Messgröße ist nur aus der Sicht eben dieser Messvorrichtung systematisch und wiederholbar. Wenn mehrere Messvorrichtungen diese identische Messgröße erfassen, so wird dieser systematische Messfehler nicht auftreten. Es werden stattdessen verschiedene Messwerte entstehen, die untereinander keinen wiederholbaren Fehler enthalten. Wenn nun mehr und mehr Messeinrichtungen (in diesem Fall Sensorknoten) Messungen vornehmen, so wird der Kommunikationsaufwand, die Messergebnisse auszutauschen, so erheblich, dass sich das System selbst überflutet. Deshalb soll diese Datenfusion der verteilten Messungen *derselben* Messgröße auf der physikalischen

Schicht der Sensorknoten mit überlagerten Funksignalen geschehen. Hierfür wird das SDJS-Protokoll herangezogen werden. So können die Knoten ihre Messergebnisse gleichzeitig aussenden und so das Skalierungsproblem in den Griff bekommen.

Die im weiteren Verlauf dargestellten technischen Lösungen werden auf der Basis von zweidimensionalen Lokationssystemen entwickelt. Sie sind jedoch ohne weiteres auf drei Dimensionen ausdehnbar.

7.5.4 Fehlerkomponenten in Messungen

Generell könnte man in einem Messsystem den Messvorgang selber als eine Operation f auf der tatsächlichen physikalischen Größe verstehen. Als Beispiel soll nun die Observation \tilde{r} einer physikalischen Größe r dienen. Die Fehler, die im Messsystem zu erwarten sind, werden zum Teil vom Systemstatus \mathbf{t} abhängen. Dieser Systemstatus kann alle Systemeigenschaften, wie z.B. die Position und Umgebungskonstellation, enthalten. Darüber hinaus wird die Messung einen rein zufälligen Fehler n tragen. Zusammenfassend kann man also die Messung folgendermaßen darstellen:

$$\tilde{r} = r + f(\mathbf{t}, r) + n. \quad (7.6)$$

Dieses Modell teilt die Einflüsse auf die Messung in die zwei Typen von Fehlern: die *systematischen* $f(\mathbf{t}, r)$ und die *statistischen*. Diese Typen sollen nun noch etwas genauer analysiert werden.

Statistische Fehler

Statistische Fehler werden durch Rauschen auf der Messeinrichtung verursacht. Typischerweise werden sie durch thermisches Rauschen verursacht. Die statistische Verteilung solcher Fehler ist typischerweise gaußisch und mittelwertfrei. Die dadurch verursachten Fehler auf dem System können als additive Messgröße auf der Observation — wie in (7.6) — dargestellt werden. Statistische Fehler können nicht vermieden werden; es gibt jedoch Maßnahmen, wie z.B. Kühlung von kritischen Bauelementen oder Filterung und Schirmung von Datenleitungen, mit denen statistische Fehler reduziert werden können. Darüber hinaus — da das statistische Modell dieser Fehler ja als gaußisch bekannt ist — werden solche Fehler

in die Schätzvorrichtungen für die physikalische Messgröße mit hineinmodelliert. Für ein Fehlermodell, das mittelwertfrei ist, wäre die einfachste Strategie, mehrere Messwerte zu mitteln, um so den Einfluss des Rauschens zu reduzieren.

Systematische Fehler

Systematische Fehler in Messsystemen hängen vom *Systemzustand* ab und können verschiedene Gründe haben. Diese reichen von Alterung über Produktionsvariation bis hin zu Dekalibrierung durch Benutzer. Auch Einflüsse der Umgebung in einer Konstellation werden als systematische Fehler verstanden. Für Lokationssysteme umfasst der Systemzustand also alle Parameter wie z.B. Orientierung und Ort der Objekte, Hardwareveränderungen sowie auch äußere Parameter wie Feuchtigkeit, Temperatur etc. Wenn die Einflüsse all dieser Parameter, die im Systemzustand \mathbf{t} zusammengefasst werden, bekannt sind, so ist die Funktion $f(\mathbf{t}, r)$ vollständig bekannt und der systematische Fehler kann aus der Observation herausgerechnet werden.

Oft sind jedoch die Einflüsse all dieser Parameter nicht vollständig bekannt oder modellierbar, doch das Messergebnis \tilde{r} wird bei unveränderten Konditionen ($\mathbf{t} = \text{const}$) wiederholbar beeinflusst. Anstatt die systematischen Fehler zu modellieren, können sie auch durch Kalibrierung erfasst werden. Methoden für Kalibrierung sind sehr unterschiedlich, aber eine genaue Kalibrierung kann ein sehr aufwendiger und kostenintensiver Prozess sein.

In Abschnitt 7.5.6 wird im Detail gezeigt, wie sich systematische Fehler in Ultraschalllokationssystemen darstellen. Es ist darüber hinaus jedoch bekannt, dass ortsabhängige, systematische Fehler auch in anderen Lokationssystemen existieren. Dies wurde z.B. für Infrarotsysteme [44] und Systeme, die mit Funkempfangsleistungsmessungen arbeiten [47], gezeigt. Generell ist also der Messvorgang oder das Medium der Messung nicht relevant. Es wird erwartet, dass das vorgeschlagene System für praktisch alle Lokalisationssysteme arbeitet.

Lokale Koordinatensysteme

In Lokationssystemen, die auf *relativen* Ortsangaben basieren, werden die relativen Positionen durch verschiedene Sicht auf das System unterschiedlich beschrieben werden. Jeder Knoten in Abbildung 7.18 (mit Großbuchstaben gekennzeichnet

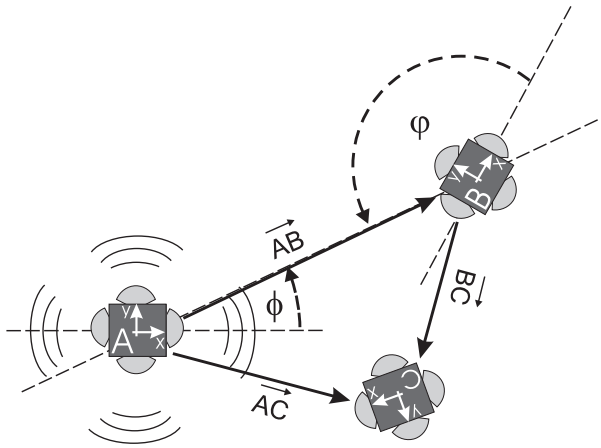


Abbildung 7.18. Relative Positionierung von Knoten

net), definiert sein eigenes, lokales Koordinatensystem, welches durch eine Vorzugsrichtung der lokalen Hardware festgelegt wird. Dies ist in Abbildung 7.18 mit den lokalen (x,y) Achsen dargestellt. Relative Positionen müssen stets mit der Angabe des Referenzsystems angegeben werden, da keine globale Orientierung oder Verankerung der Darstellung existiert. Deshalb müssen relative Koordinaten mit der zusätzlichen Angabe des Referenzsystems angegeben werden:

$$\overrightarrow{AB}_A$$

Die Notation (7.7) gibt ein Beispiel für die Angabe einer relativen Position. Sie drückt die Position von B (also den Vektor von A nach B) aus der Sicht des Referenzsystems von A aus. Wenn relative Positionen über mehrere Schritte zusammengesetzt werden sollen, so ist es nötig, alle Koordinaten in dasselbe Referenzsystem abzubilden. Im kartesischen Fall kann man dann die Vektoren einfach komponentenweise addieren. Die Position von Knoten C in Abbildung 7.18 aus

der Sicht von A ($=\vec{AC}_A$) kann also Stück für Stück als Vektorkette von A nach C zusammengesetzt werden. Dies würde wie folgt aussehen:

$$\begin{aligned} AC_A &= \vec{AB}_A + \vec{BC}_A \\ \vec{BC}_A &= \begin{pmatrix} \cos \rho & -\sin \rho \\ \sin \rho & +\cos \rho \end{pmatrix} \vec{BC}_B \\ \rho &= \Phi + (180^\circ - \varphi) \end{aligned} \quad (7.7)$$

Wichtig zu bemerken ist hier, dass der Sendewinkel ϕ und der Empfangswinkel φ zwischen zwei Knoten austauschbar sind: Der Sendewinkel von A nach B ist zugleich der Empfangswinkel von A, wenn B sendet. Die Fehlermodelle und Statistiken von den gemessenen Sende- und Empfangswinkeln können sich jedoch deutlich unterscheiden.

7.5.5 Die Plattform der Experimente

Die Sensorknoten, die als Studienobjekte innerhalb des RELATE-Projektes [5] implementiert wurden, können Entfernungen und Winkel messen. Abbildung 7.19 zeigt einen solchen Prototypen mit vier Ultraschallkomponenten, die sowohl als Sender als auch als Empfänger verwendet werden. Im weiteren Verlauf wird er auch „Relateobjekt“ genannt. Mit dieser Ultraschallplattform wurden umfangreiche Experimente durchgeführt. Fünf solcher Prototypen wurden auf einem 1.6 mal 2.4 m Tisch in 50 verschiedenen räumlichen Konstellationen aufgebaut und dort jeweils für 7 Minuten tausende von Messungen genommen. Die Genauigkeit der rohen Messungen von Entfernungen und Empfangswinkeln ist in Abbildung 7.20 zu sehen.

Bestimmung des Empfangswinkels

Ein Sensorknoten, der eine Messung durchführt, kann normalerweise auf mehreren Ultraschallwandlern ein erkennbares Signal aufgreifen. Da die Schallwandler eine winkelabhängige Sensitivität haben, wird eine Empfindlichkeitsverteilung ähnlich der in Abbildung 7.21 erwartet. Diese Abbildung zeigt die normalisierten Empfindlichkeitskurven von drei Ultraschallwandlern eines Relateobjektes. Zur Berechnung des Empfangswinkels wurden die Empfindlichkeitskurven durch ein

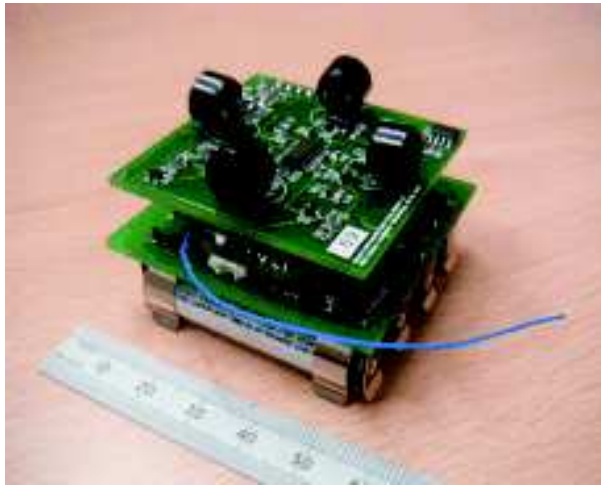


Abbildung 7.19. Ein „Relateobject“ ist ein mobiler Sensorknoten für Ultraschallmessungen

lineares Modell approximiert. Dies vereinfacht die Berechnung auf den Mikroprozessoren erheblich. Wenn man nun die einkommenden Empfangsleistungen von drei Schallwandlern gemessen hat, kann man den Empfangswinkel bestimmen. Ähnliche Modelle, die Intensitätsverteilungen von benachbarten Schallsensoren verwenden, wurden z.B. auf Roboterplattformen vorgeschlagen [64].

Für die Lokalisation von Schallquellen mit Hilfe von Mikrophonarrays [6] ist es verbreitet, anstelle von Intensitäten, die unterschiedlichen Ankunftszeiten zu verwenden, um die Position der Quelle zu schätzen. Für die kompakte Hardware (Abbildung 7.19) aus dem RELATE-Projekt ist es jedoch nicht möglich, ein Verfahren zu verwenden, das die Ankunftszeiten der Schallwellen zwischen den Schallwandlern als Schätzbasis verwendet. Der Abstand zwischen den Schallwandlern ist typischerweise 3cm, was einer unterschiedlichen Ankunftszeit von $[-100\mu s..100\mu s]$ entspricht, wenn das Objekt über 180° gedreht wird. Um Zeiten im Bereich von μs auflösen zu können, sind Abtastraten deutlich über 100kHz und ein leistungsfähiger DSP notwendig. In [53] wird z.B. ein Lokalisationssystem mit einem Mikrophonarray vorgestellt, das Laufzeitdifferenzen verarbeitet. Es wird dort eine 5MHz-Taktrate zur Zeitbestimmung verwendet. Andere Systeme

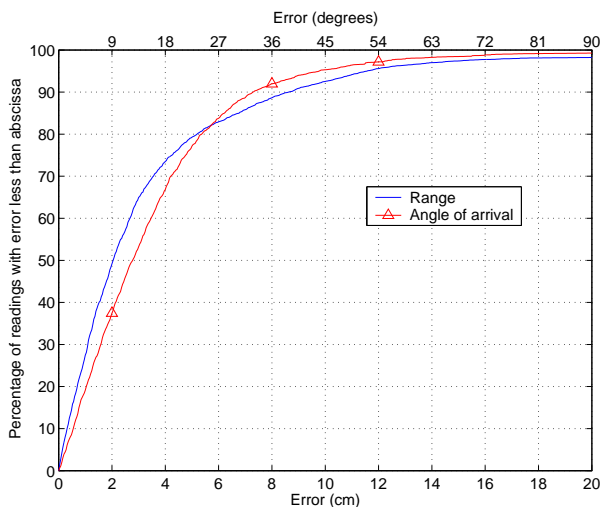


Abbildung 7.20. Genauigkeit der rohen Messwerte für die Ultraschall-Plattform aus [20, Abbildung 11]

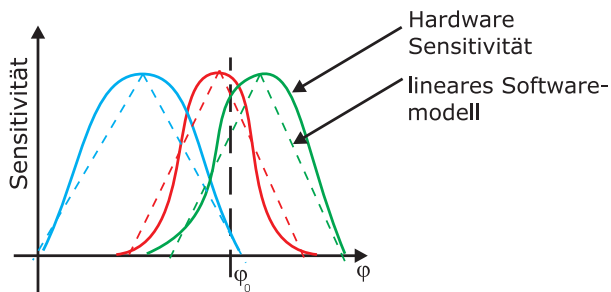


Abbildung 7.21. Verteilung der Empfindlichkeit der Schallwandler in Abhängigkeit vom Einfallswinkel φ

me, die Laufzeitdifferenzen verwenden, haben typischerweise viel größere Abstände zwischen Schallwandlern, z.B. 50cm in [79].

Aufgrund dieser technischen Einschränkungen wurde statt Laufzeitdifferenz die Empfangsamplitude verwendet, um den Empfangswinkel zu bestimmen. Ähnliche Verfahren wurden in [7], [60] dargestellt. Diese Veröffentlichungen basieren zwar auf hörbaren Audiosignalen; die Theorie ist aber dennoch relevant und in vielen Fällen sogar direkt auch auf Ultraschallsignale anwendbar. Die Autoren modellieren die Sensitivität und das Emissionsverhalten als Kosinusfunktion vom aus- oder eintretenden Winkel. So konnte z.B. die Position und die Kopfausrichtung von sprechenden Personen in einem Raum festgestellt werden, indem eine große Anzahl von Mikrofonen an den Wänden angebracht wurde. Wie bereits erwähnt, wurde für die vorliegende Arbeit aus Implementierungsgründen ein lineares Modell (also kein Kosinus) verwendet. Darüber hinaus mussten wegen der geringen Abtastrate von nur 30kHz und nur 4 Ultraschallwandlern einige Heuristiken angewandt werden. Als Basis der Detektion diente z.B. nur die Einhüllende des Ultraschallsignals, da 30kHz Abtastrate nicht genug für die Abtastung der Trägerfrequenz des Ultraschalls (ca. 40kHz) ist.

7.5.6 Empirische Fehlermodelle

In diesem Abschnitt werden nun die Fehler untersucht, die während der Experimente auftraten. Die Analysen beziehen sich teilweise auf ein oder mehrere Objekte.

Fehler in der Bestimmung der Empfangswinkel

Abbildung 7.22 zeigt ein Beispiel einer Messfehlerkurve für einen Ultraschallprototypen. Sie zeigt den absoluten Fehler des gemessenen Empfangswinkels (y-Achse) gegen den tatsächlichen Empfangswinkel (x-Achse). In manchen Winkelbereichen scheint der Fehler eher zufällig (z.B. der Bereich $[-40^\circ.. -20^\circ]$), was nicht der Annahme entspricht, dass benachbarte Schallwandler einen fließenden Übergang erzeugen. Deshalb wurden die Fehler einer genaueren Analyse unterzogen. Basierend auf (7.6) wird nun die Messung des Empfangswinkels wie folgt modelliert:

$$\tilde{\varphi} = \varphi + f(\mathbf{t}) + n \quad ; \mathbf{t} = (\varphi, \Phi, r) \quad (7.8)$$

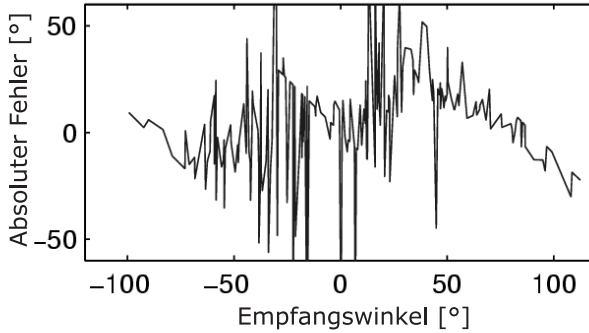


Abbildung 7.22. Messfehlerkurve für den Empfangswinkel

Anschließend wurden Möglichkeiten gesucht, *systematische* und *statistische* Fehler zu charakterisieren und voneinander zu separieren. Nimmt man den Mittelwert des Fehlers von wiederholten Messungen (gemeint ist hier, dass der Systemstatus währenddessen konstant ist; $\mathbf{t} = \text{const}$), so erhält man:

$$E(\tilde{\varphi} - \varphi) = E(f(\mathbf{t})) + E(n) = f(\mathbf{t}) + 0 = f(\mathbf{t}) \quad (7.9)$$

Damit kann man den systematischen Teil der Fehler aus den Daten extrahieren. Kennt man nun den systematischen Fehler $f(\mathbf{t})$ und den tatsächlichen, richtigen Wert, kann man den statistischen Fehler n durch einfache Subtraktion „(7.9) - (7.8)“ isolieren. Mit dieser einfachen Analysevorschrift wurden die Daten untersucht. Der systematische Fehler wird in Abbildung 7.23 auf verschiedenen Entfernungen dargestellt. Da es zwischen nahen und weiträumigen Experimenten keinen sichtbaren Unterschied gibt, wurde dieser Fehler mit einer Gaußverteilung ($\mu = 0; \sigma = 14^\circ$) angenähert. Der Systemzustand \mathbf{t} wird zu $\mathbf{t} = (\Phi, \varphi)$ kontrahiert, wobei die Abhängigkeit vom Sendewinkel auch noch vernachlässigt wird, da alle 4 Schallwandler zur gleichen Zeit angeregt werden können. Der Zustand \mathbf{t} vereinfacht sich also zu $\mathbf{t} = \varphi$ und (7.8) wird:

$$\tilde{\varphi} = \varphi + f(\varphi) + n \quad (7.10)$$

Um zu zeigen, dass die Modelle (7.10) und (7.9) tatsächlich auf die Messwerte passen, soll nun gezeigt werden, dass systematische Fehler bei konstantem Systemzustand wirklich wiederholbar sind. In Abbildung 7.24 sieht man, wie die

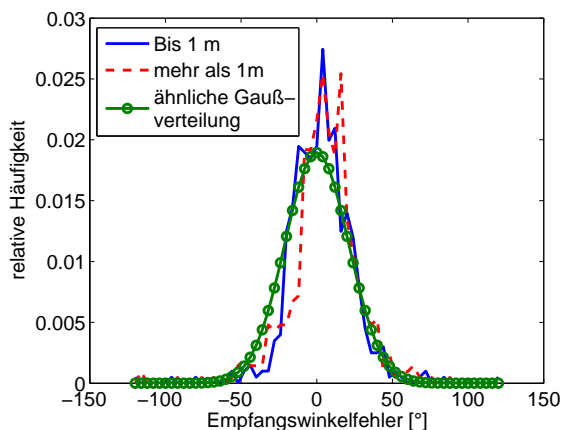


Abbildung 7.23. Mittlerer Fehler für den Empfangswinkel bei verschiedenen Entfernungen

Fehler auf verschiedenen Winkeln zwar unterschiedlich sind, aber wiederholbar auf den einzelnen Winkeln. Die Streuung bei festgehaltenem Empfangswinkel ist relativ gering. Der statistische Anteil ist in Abbildung 7.25 zu sehen. Er ist quantitativ deutlich kleiner als der systematische, was die Motivation weiter bestärkt, sich in besonderem Maße der systematischen Fehler anzunehmen.

Für die weiteren Betrachtungen wird der nun der systematische Teil des Fehler des Empfangswinkels wie in Abbildung 7.23 modelliert; der statistische Fehler als gaußisch ($\mu = 0; \sigma = 1.4^\circ$) wie in Abbildung 7.25. Dieses detaillierte, empirische Modell ist auch die Basis für die Simulationen in Abschnitt 7.5.11. Diese genaue Modellierung gibt eine wesentlich bessere Beschreibung der tatsächlichen Hardware als es reine Fehlergraphen im Stil von Abbildung 7.22 erreichen.

Fehler in der Bestimmung der Sendewinkel

Der Sendewinkel kann bestimmt werden, in dem die vier Schallwandler nicht gleichzeitig, sondern nacheinander angesprochen werden. Der Empfänger nimmt dann den Schallwandler, der die stärksten Empfangsleistungen verursacht hat, als

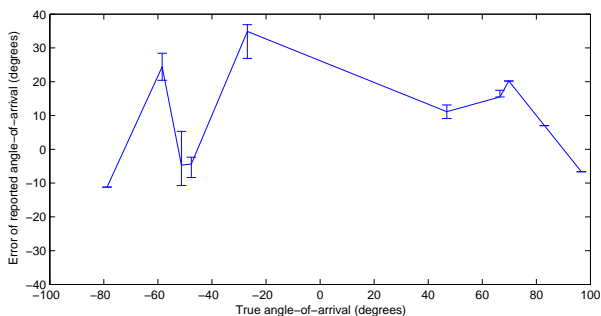


Abbildung 7.24. Wiederholen von Empfangswinkelmessungen bei gleichem Systemzustand t . Die Fehlermaße repräsentieren die 5%- und 95%-Bereiche. Die Streuung ist relativ klein — dies zeigt, dass die Fehler tatsächlich wiederholbar identisch, also systematischer Natur sind.

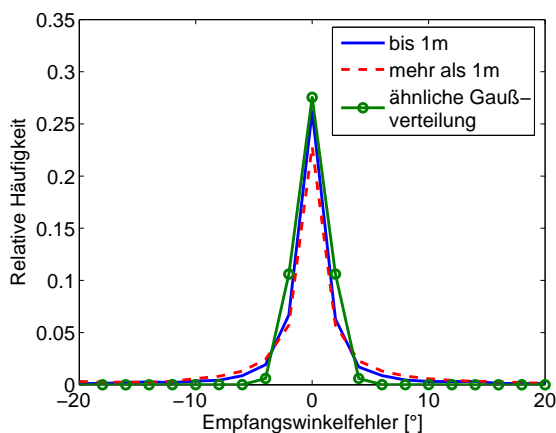


Abbildung 7.25. Die Wahrscheinlichkeitsdichte für das Rauschen auf dem Empfangswinkel

Sendewinkel an. Mit vier Schallwandlern, die jeweils 90° abdecken, wird so für den Fehler eine Gleichverteilung zwischen $[-45^\circ..45^\circ]$ entstehen.

Fehler in der Abstandsbestimmung

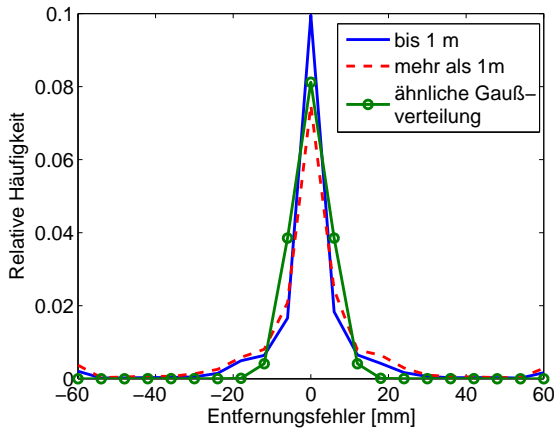


Abbildung 7.26. Der statistische Fehler in der Abstandsmessung, verursacht durch Quantisierung und Rauschen

Auch der Fehler in der Abstandsbestimmung kann in *systematische* und *statistische* Teile zerlegt werden. Der systematische Teil hängt hauptsächlich von der Temperatur des Mediums ab, da die Schallgeschwindigkeit c von dieser abhängt:

$$c = \sqrt{\kappa R T} \quad (7.11)$$

mit $R = 287.05 \frac{J}{kg \cdot K}$ für Luft und T als absolute Temperatur in Kelvin. Mithilfe einer zusätzlichen, lokalen Messung, hat man eine gute Schätzung für die Temperatur der Umgebung und kann diesen Einfluss herausrechnen. Der statistische Teil der Fehler wird hauptsächlich durch Quantisierung und thermisches Rauschen verursacht. In Abbildung 7.26 sieht man seine Verteilung. Diese Kurve wurde auch aus den Messwerten der Experimente extrahiert. Der mittlere Fehler der Entfernung (also der systematische Teil) ist in Abbildung 7.27 zu sehen. Es

hat sich gezeigt, dass hier keine Abhängigkeit vom Sende- oder Empfangswinkel besteht. Dieser Fehler zeigt jedoch ein sehr unsymmetrisches Verhalten. Das

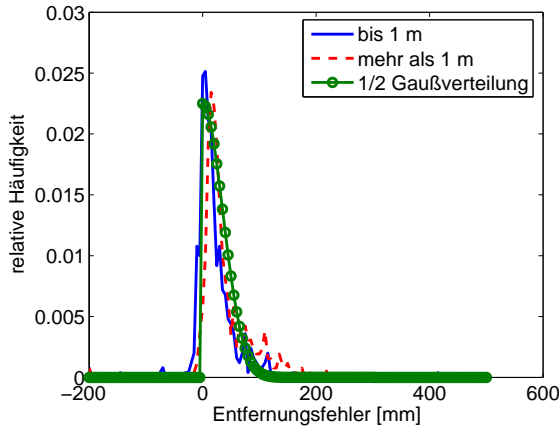


Abbildung 7.27. Der mittlere Fehler der Abstandsmessung

Ultraschallsystem tendiert dazu, „zu weit“ zu messen. Die Verteilung ist nicht mittelwertfrei und auch nicht symmetrisch.

Für die weiteren Betrachtungen wird der statistische Fehler wieder als gaußisch mit $\mu = 0$; $\sigma = 4.9\text{mm}$ modelliert; der systematische Teil wird als halbseitige Gaußverteilung mit $\mu = 0$; $\sigma = 3.5\text{cm}$ modelliert. Dieses Modell ist wiederum auch Basis für die Simulationen.

7.5.7 Messen und Mittelung von Messwerten in Lokationssystemen

Die größten *systematischen Fehler* in den Messungen sind für den Empfangswinkel gefunden worden. Neben Kalibrierung — die in diesem Zusammenhang nicht betrachtet werden soll — wäre es ein naheliegender Ansatz, den Empfangswinkel zu variieren, um so mehrere Messungen unter verschiedenen Systemzuständen \mathbf{t} zu bekommen. Hierfür müssten die Knoten aktiv rotiert werden. So bekäme man mehrere Stichproben der Verteilung aus 7.23. Mit diesen mehreren Stichproben könnte man nun den Empfangswinkel besser bestimmen. Durch die Rotation der

Objekte würde der Systemstatus zwischen den Stichproben nicht konstant bleiben, und es entsteht kein systematischer, wiederholbarer Fehler mehr. Der Gewinn einer solchen Vorgehensweise lässt sich sogar quantisieren:

Mathematischer Exkurs: Mittelung von Messwerten

Gegeben sei ein Zufallsprozess einer Messung. Er soll hier mit der Variablen X ($\mu_X; \sigma_X^2$) beschrieben werden. Nach der oben gezeigten Fehleranalyse wird erwartet, dass die Verteilung mittelwertfrei ($\mu_X = 0$) ist. Der Mittelungsprozess erzeugt nun eine neue Variable, die mit Y bezeichnet wird und sich zu

$$Y = \frac{1}{N} \sum_{i=1}^N X \quad (7.12)$$

ergibt. Die Momente sind entsprechend $\mu_Y = 0$ und $\sigma_Y^2 = \frac{\mu_X}{N}$. So wird also eine schmalere Verteilung erzeugt, deren Varianz gegenüber X um N skaliert ist. Dies ist der natürliche Effekt der Mittelung. Damit dies auch für die systematischen Fehler funktioniert, muss dafür gesorgt werden, dass die N Stichproben für die Mittelung tatsächlich unabhängig voneinander gezogen werden und nicht auf einen wiederholbaren Fehler führen. Dafür muss also der Systemzustand variiert werden.

Mit einem bekannten N und σ_E^2 kann man nun die Wahrscheinlichkeit für bestimmte Fehler berechnen. Genauso lässt sich auch die Verbesserung durch die Mittelung quantifizieren. In Abbildung 7.28 sieht man die kumulative Verteilung der Fehler. Die niedrigste Kurve ist diejenige ohne Mittelung, also die der einfachen Messwerte (der Messwertefehler wurde für diese Darstellung als gleichverteilt zwischen 0 und 50cm angenommen). Eine weitere, interessante Veranschaulichung bietet Abbildung 7.29. Dort wird die 95% Grenze für die Messfehler gezeigt. Selbst, wenn man mit einer 95% Fehlergrenze von 50cm startet, kann man nur durch Mittelung von nur 30 Messwerten, die Grenze von 10cm mit einer Zuverlässigkeit von 95% unterschreiten.

Mit diesem Exkurs sollte quantitativ gezeigt werden, wie Mittelungen den Messprozess verbessern können. Nun wird im Detail dargelegt, wie dies auf das Szenario von Ultraschalllokalisationssystemen angewendet werden kann.

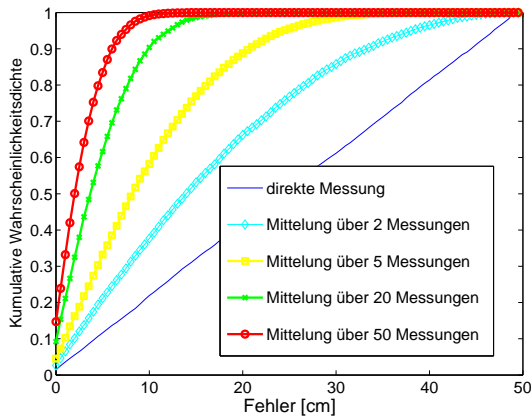


Abbildung 7.28. Die Verbesserungen in der Genauigkeit, wenn mehrere Messwerte gemittelt werden; die Abbildung zeigt kumulative Wahrscheinlichkeitsdichten für verschiedene Anzahlen von Messungen, die gemittelt werden.

7.5.8 Mittelung über mehrere Messpfade

In Abbildung 7.30 sieht man ein typisches Szenario für relative Positionierung und Lokalisation ohne Infrastruktur. Die Knoten (benannt A bis F) lokalisieren sich gegenseitig, indem sie Winkel und Entfernungen zueinander bestimmen. Hier soll auch wieder angenommen werden, dass es sich um ein Ultraschallsystem handelt. Im Beispiel von Abbildung 7.30 will Knoten A die relative Position von Knoten B bestimmen.⁵ Die Fehler, die hierbei entstehen, besitzen komplexe Abhängigkeiten mit dem Systemzustand, wie bereits in Abschnitt 7.5.6 ausgeführt. Wenn man typische Fehlerkurven wie z.B. in Abbildung 7.22 zu Grunde legt, so wird man einen erheblichen *systematischen* Fehler für die relative Position zwischen A und B bekommen. Dieser wird hauptsächlich von der Winkolorientierung von A und B abhängen. In dem mathematischen Exkurs zur Mittelung wurde gezeigt, dass Mittelung die gesamte Genauigkeit erheblich verbessern kann. Um

⁵Die relative Position hängt auch vom Koordinatensystem ab. In diesem Fall würde A das Referenzsystem definieren. Mit Hilfe von Laufzeit und Amplitudenmessungen und den Techniken aus Kapitel 7.5.4 und 7.5.5 können die relativen Positionen der Knoten untereinander bestimmt werden.

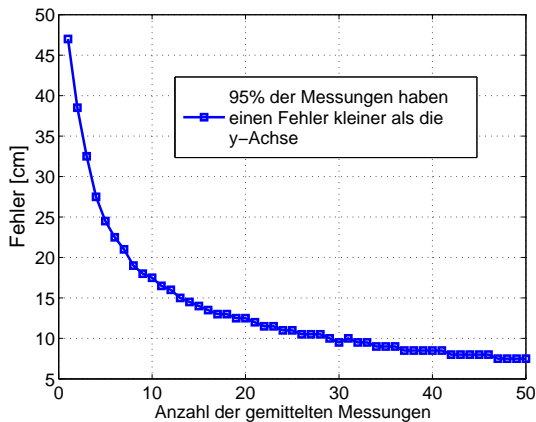


Abbildung 7.29. Die 95% Fehlerkurve in Abhängigkeit von der Anzahl der gemittelten Messwerte: die Mittelung von nur wenigen Messwerten erzeugt schon eine relevante Verbesserung.

aber aus der Mittelung Vorteil ziehen zu können, ist es notwendig, Abstände und Winkel unter *verschiedenen* Systemzuständen zu messen. Die Abhängigkeit des Fehlers von der Orientierung würde hier also bedeuten, dass die Knoten A und B zwischen Messungen *rotiert* werden müssten. Diese Messungen könnten dann gemittelt werden, um die Verbesserung zu erreichen. Obwohl dieser Prozess die Ortsbestimmung zwischen A und B sicherlich verbessern würde, ist er unpraktisch, ja unmöglich aufgrund folgender Überlegungen:

- Jede Messung, die in den Mittelungsprozess einfließen soll, verursacht einen Prozessaufwand von Synchronisation, Laufzeitmessung etc. Für N Messungen müssen N Messprozesse ausgeführt werden, die Skalierung ist $O(N)$.
- Objekte zu rotieren ist unpraktisch oder sogar unmöglich. Die Knoten, die an der Lokalisierung teilnehmen, könnten z.B. Alltagsgegenstände sein, die mit Elektronik versehen wurden. Deshalb kann es nicht vorausgesetzt werden, dass alle Objekte sich aktiv bewegen, also rotieren können.

Dies sind nur zwei wichtige Gründe, die es verhindern, die Idee der rotierenden Objekte zur Verbesserung der Messung weiter zu verfolgen. Das System skaliert

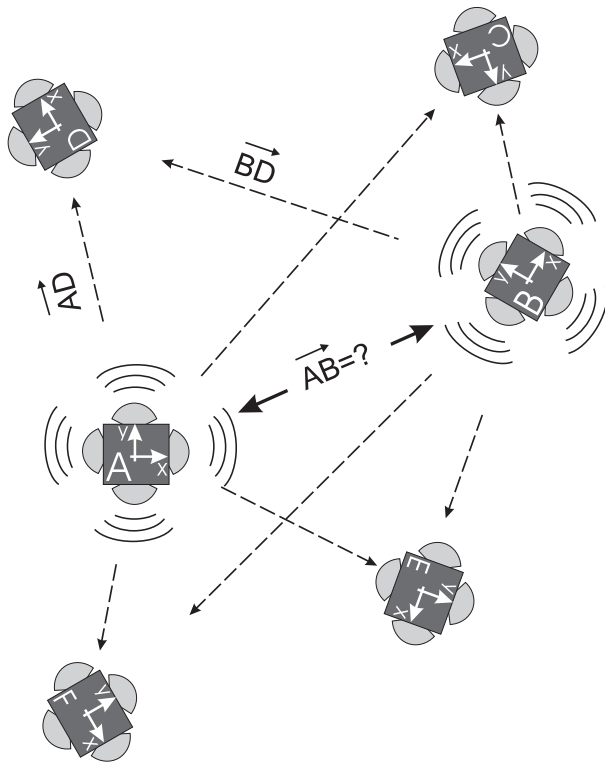


Abbildung 7.30. Mehrere Messpfade führen von Knoten A nach B

schlecht und ist schlicht mechanisch nicht durchführbar. Deshalb wird nun ein alternativer Weg vorgeschlagen, wie man mehrere Messungen erzeugt, ohne z.B. die Objekte bewegen zu müssen.

Mehrfaches paralleles Mitteln

In dem Referenzszenario aus Abbildung 7.30 kann man zwei wichtige Eigenschaften des Systems feststellen:

1. Die Ultraschalltechnologie hat eine Rundrufnatur. Dies gilt auch für andere Medien wie Infrarot, Funkwellen und Audio.
2. Es wird eine Umgebung angenommen, in der viele Objekte vorhanden sind, die potentiell an Lokalisierungen und Messungen teilnehmen können.

Ausgehend von diesen Annahmen, gibt es nun eine Möglichkeit, mehrere Messungen zur gleichen Zeit zu nehmen: *immer, wenn ein Knoten ein Signal für Messungen aussendet, kann jeder andere Knoten in der Umgebung dieses Signal benutzen, um eine Messung durchzuführen.* Dies ist ein interessanter Effekt, der sich auf die Skalierbarkeit des Systems auswirkt. Wenn N Knoten in Reichweite des Messsignals sind (z.B. ein Ultraschallimpuls), können sie alle zur gleichen Zeit Messungen vornehmen. Damit wird die Skalierung dieser N Messungen in der Laufzeit von $O(N)$ auf $O(1)$ verbessert.

Anhand von Knoten D soll der genaue Ablauf im Referenzbeispiel erklärt werden: Wenn A ein Ultraschallsignal aussendet, kann D seinen Empfangswinkel und die Entfernung zu A, und so letztlich die *relative Position* von A bestimmen. Wenn B Ultraschallsignale aussendet, kann D ebenfalls die relative Position zu B bestimmen. Wenn D darüber hinaus den Sendewinkel von A kennt, so kann es sein lokales Koordinatensystem in die Sicht von A drehen (siehe (7.7)) und kann so die relative Position von A zu B durch einfache Vektoraddition bestimmen:

$$\overrightarrow{AB}_A = (\overrightarrow{AD}_D - \overrightarrow{BD}_D)_A \quad (7.13)$$

Alle anderen Knoten in der Umgebung (C, E, F) können dies ebenfalls zur gleichen Zeit tun, und werden anschließend alle die relative Position von B aus der Sicht von A (\overrightarrow{AB}_A) bestimmen können. Nachdem also A und B jeweils einmal einen Ultraschallimpuls ausgesendet haben, sind N Messwerte der relativen Position von A nach B im System vorhanden. Problematisch ist nur, dass diese N Messwerte *verteilt* vorliegen, also als lokales Ergebnis in den Knoten C, D, E, F.

Es ist also ein weiterer Kommunikationsschritt ($O(N)!$) nötig, um die N Messwerte einzusammeln und zu mitteln.

Um die Messwerte zu mitteln und daraus den Gewinn zu erzielen, wie es die Mittelung verspricht, muss man sicherstellen, dass die Voraussetzungen für diesen Prozess auch erfüllt sind. Der Mittelungsprozess in Abschnitt 7.5.7 nahm an, dass ein Knoten rotiert würde, um so mehrere Stichproben unter verschiedenen Systemzuständen (also verschiedenen Orientierungen) zu bekommen. Im Prozess in Abbildung 7.30 werden verschiedene, zufällige Empfangswinkel in den verteilten Knoten C, D, E und F auftreten und so die Stichproben unter verschiedenen Orientierungen genommen. Die zufällige Verteilung im Raum sorgt für „genügende“ Unabhängigkeit der Stichproben.

7.5.9 Datenfusion mit Multi-SDJS

Im vorangegangenen Abschnitt wurde gezeigt, wie es möglich ist, mehrere Messungen derselben Messgröße effizient zu erzeugen. Dies war hauptsächlich auf die Rundrufnatur der Signale für den Messvorgang zurückzuführen. Das einzige Problem ist nun noch, dass die Messergebnisse nach dem Messprozess in den Knoten verteilt sind und erst wieder zusammengeführt werden müssen. Für die Sammlung der lokalen, verteilten Ergebnisse wird wieder SDJS vorgeschlagen. So können die Messergebnisse effizient und mit konstanter Skalierung in der Prozesszeit eingesammelt werden, und der „Masseneffekt“ der verteilten Objekte kann genutzt werden.

Der Ausgangspunkt ist nun, dass N Messwerte \tilde{r}_i des gleichen Vektors \vec{AB}_A vorliegen, die eingesammelt und gemittelt werden sollen. Hierfür wird der Multi-SDJS-Prozess eingesetzt werden, wie er schon im Supermarkt-Szenario in Abschnitt 7.3.3 und für Temperaturwerte in Abschnitt 7.4 verwendet wurde.

Da der Multi-SDJS-Prozess nur auf skalaren Größen arbeiten kann, werden die Messwerte in ihre Komponenten $(\tilde{x}_i, \tilde{y}_i)^T$ oder $(\tilde{\varphi}, \tilde{r})^T$ zerlegt. Dann kann ein Multi-SDJS-Prozess auf jeder dieser Komponenten durchgeführt werden. Für den SDJS-Prozess wird eine Vollvermaschung in der Netzwerktopologie der Knoten angenommen. Dies wird sicherlich für die meisten Fälle zutreffen, da die Reichweite der Funkverbindung die Reichweite von den Ultraschallsignalen sicherlich übertreffen wird. Abbildung 7.31 zeigt, wie das Einsammeln der Daten funktioniert: Knoten A initialisiert den Multi-SDJS-Prozess mit dem Aussenden eines

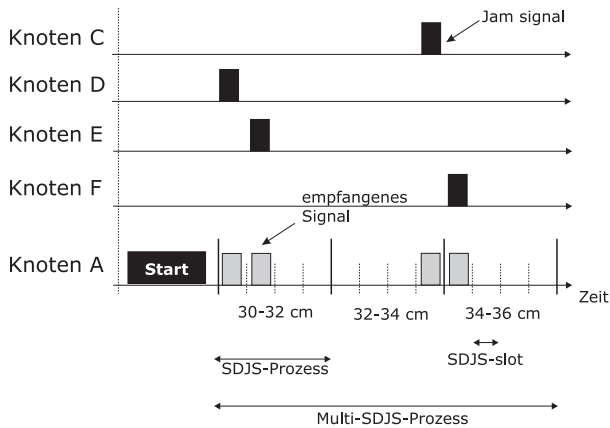


Abbildung 7.31. Verteilte Messwerte werden mit dem Multi-SDJS-Prozess eingesammelt

Startpaketes. In diesem Fall enthält das Paket die Information, dass drei SDJS-Prozesse folgen, mit je vier SDJS-Slots. Die zugehörigen Messwerte sind 30-32cm, 32-34cm und 34-36cm. Nach dem Startpaket verfahren die Knoten wie immer: sie wählen den richtigen SDJS-Prozess aus und wählen zufällig einen Zeitslot. Basierend auf den Jamsignalen in den Zeitschlitzen, können die Anzahlen dann geschätzt werden, und die Knoten können so die Messwerte kommunizieren. Da für das Lokationssystem besonders die Mittelung der Werte interessant ist, können die Ergebnisse von Abschnitt 7.4 hier als direkter Nachweis des Konzeptes dienen.

7.5.10 Der Ablauf der Lokalisierung

Für die Zusammenfassung soll noch einmal das Szenario in Abbildung 7.30 betrachtet werden. Es wird angenommen, dass am Anfang des Prozesses keine Messungen oder weitere Informationen vorliegen. Knoten A will die relative Position von B \vec{AB}_A bestimmen. Die folgenden Schritte werden dafür abgearbeitet:

- 1) **Start:** eine neue Position (z.B. \vec{AB}_A) soll festgestellt werden; dies wird durch eine Funknachricht mitgeteilt.

- 2) **Physikalische Messung 1:** Knoten, die Endpunkte der Messung sind (hier: A und B), senden Signale für die Messungen aus. Skalierung in Prozesszeit: $O(1)$.
- 3) **Physikalische Messung 2:** Alle anderen Knoten in der Umgebung verwenden die gesendeten Signale, um damit Messungen von Winkeln und Entfernungen zu gewinnen. Skalierung in Prozesszeit: $O(1)$.
- 4) **Messungsaufbereitung:** Die lokalen Messungen werden in das Referenzkoordinatensystem transformiert, um eine konsistente Sicht zu erzeugen. Skalierung in Prozesszeit: $O(1)$.
- 5) **Sammeln der Ergebnisse:** Einer der Endpunkte (A oder B) startet den Multi-SDJS-Prozess; alle Knoten quantisieren ihre Ergebnisse und nehmen am Multi-SDJS-Prozess teil. Skalierung in Prozesszeit: $O(1)$.
- 6) **Datenfusion:** Die Endpunkte verwenden die SDJS-Schätzungen als Basis und berechnen den Mittelwert der gesammelten Ergebnisse. Skalierung in Prozesszeit: $O(1)$.
- 7) **Ende:** Die neue Position \vec{AB}_A ist bestimmt und kann per Rundruf verteilt werden.

7.5.11 Anwendung und Simulation

In diesem Abschnitt werden Szenarien vorgestellt, die mit OMNet++⁶ simuliert wurden. In der Simulation wurde Szenarien in einem Raum der Größe 5x5m betrachtet. Die Anzahl der Knoten reichte bis zu 200. Die Fehlermodelle wurden aus der Analyse der Experimente (siehe Abschnitt 7.5.6) mit der Ultraschallplattform eingebracht. Zusammengefasst gilt für die Fehlermodelle:

Fehlermodell für die Empfangswinkel: der systematische und statistische Teil sind beide als Gaußdichte modelliert.

Fehlermodell für den Sendewinkel: Gleichverteilung über $[-45^\circ \dots +45^\circ]$.

Fehlermodell für die Abstandsmessung: Hier wurde die rechtsseitige Gaußverteilung verwendet sowie eine weitere Gaußverteilung für das statistische Rauschen. In der Simulation wurde für den systematischen Fehler statt der rechtsseitigen auch eine symmetrische Gaußverteilung ausprobiert. Die Ergebnisse unterschieden sich nicht relevant. Dies liegt daran, dass der Fehler in der Entfernungsbestimmung relativ klein im Vergleich zu den Winkelfehlern ist.

⁶<http://www.omnetpp.org>

In den Simulationen wurden *direkte Messungen* mit den Ergebnissen des Mittelungsprozesses verglichen. Direkte Messungen sind solche, die nur auf Messungen der Endpunkte beruhen, ohne weitere „Helfer“⁷ mit einzubeziehen. Hierbei wurde stets das bestmögliche Ergebnis verwendet, also A und B in der Signalisierung vertauscht, wenn dies bessere Ergebnisse erzielt hat. Dies macht Sinn, denn die Modelle für Empfangs- und Sendewinkel sind ja nicht gleich! Wenn also die Position von B aus der Sicht von A bestimmt werden sollte, so wurden hierfür in der Simulation sowohl Signale von A als auch von B ausgesendet. Die Ergebnisse der gegenseitigen Positionierung wurden verglichen, und das bessere Ergebnis als die *direkte Messung* bezeichnet.

Relative Position zwischen zwei Knoten

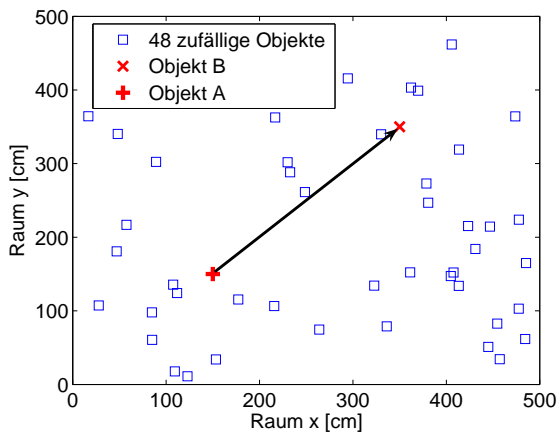


Abbildung 7.32. Ein Beispiel aus der Simulation für eine Konstellation mit 50 Knoten

Abbildung 7.32 zeigt einen simulierten Raum mit 50 Knoten. Es wird die relative Lokation von A (150, 150) zu B (350, 350) bestimmt. Es wurden 1000 Simulationsläufe mit verschiedenen Positionen der 48 Helferknoten durchgeführt.

⁷Als Helfer werden diejenigen Knoten oder Relateobjekte bezeichnet, die nicht Endpunkte in dem Messvorgang sind, aber die Ortung der Endpunkte mit zusätzlichen Messungen unterstützen.

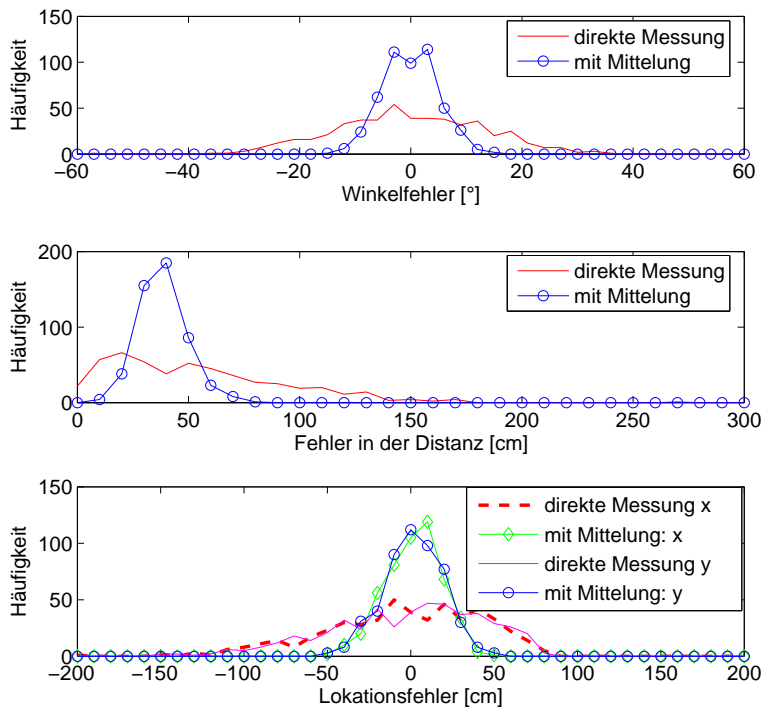


Abbildung 7.33. Die Messungen können durch Mittelung verbessert werden: ihre Varianz wird kleiner, die Dichtekurven werden zusammengeschoben

Die systematischen Fehlermodelle wurden dabei pro Lauf und Knoten neu erzeugt. In Abbildung 7.33 werden die Ergebnisse der direkten Messung mit den gemittelten verglichen. Die drei Graphen zeigen die relativen Häufigkeiten der absoluten Fehlerwerte für den Winkel Φ , den Abstand von \hat{B} zu B sowie die ermittelten kartesischen Positionen von B im Referenzsystem des Raumes. Alle Graphen vergleichen direkte Messungen mit den gemittelten.

Für die direkten Messungen entsteht der maximale Fehler in dem Fall, wenn Knoten B (fälschlicherweise) auf der gleichen y-Position wie A lokalisiert wird. Dann ist die x-Position von B 432cm plus Rauschen; somit ist also der Fehler in x-Richtung 82cm. Größere Fehler in positiver x-Richtung können nicht auftreten. Man sieht diese Grenze in der x-Position in Abbildung 7.33. Wie zu erwarten, sind auch die Fehler der Winkel symmetrisch zu Null. Betrachtet man die Verteilung der Winkelfehler für die direkten Messungen, so kann man die Verteilung aus Abbildung 7.23 wiedererkennen. Sie ist Basis der Simulation. Für alle Messgrößen kann man erkennen, wie die Mittelungen die Ergebnisse verbessern. Die Verteilungen sind erheblich schmäler. Wichtig ist noch zu erwähnen, dass unter „Abstand“ hier der Abstand der geschätzten Position von B zu seiner wirklichen verstanden wird — nicht der Abstand zwischen A und B! Abbildung 7.34 zeigt ei-

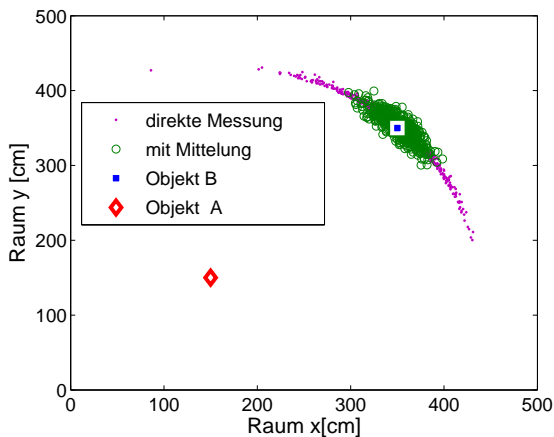


Abbildung 7.34. 1000 Simulationen mit 50 Knoten

ne weitere Illustration von 1000 Simulationsläufen mit verschiedenen Positionen der Knoten. Die Positionen von Knoten A und B wurden dabei nicht verändert. Die Abbildung zeigt, wie die direkten Messungen sich auf einem Kreissegment verteilen. Dies liegt daran, dass der Winkelfehler viel mehr streut als die Entfernungsmessung. Nach dem Mittelungsprozess mit den 48 Hilfsknoten sind die Ergebnisse stärker um den Zielpunkt B konzentriert.

(μ/σ^2)	direkte Mess.	mit Mittelung (N=48)
Φ [°]	(-0.3/ 180.0)	(-0.1/ 24.2)
abs. [cm]	(53.6/ 1502)	(38.5/ 116.1)
x [cm]	(-4.7/ 2232)	(1.4/ 296.2)
y [cm]	(-2.6/ 2121)	(1.6/ 313.0)

Tabelle 7.1. Statistische Charakteristika für N=0, 48

In Tabelle 7.1 sind die statistischen Kenngrößen für die Verteilungen der direkten und gemittelten Messungen zu finden. Die Verbesserungen sind deutlich in der Verringerung der Varianzen sichtbar. Durch die Mittelung erreicht man deutlich schmalere Verteilungen. Mit (x,y) wird der Positionsfehler von B im Raumreferenzsystem dargestellt. Φ bezeichnet den Winkel, unter dem der Knoten A den Knoten B gesehen hat. Mit dem absoluten Fehler wird der absolute Positionsfehler von B zu seiner tatsächlichen Position (350,350) bezeichnet. In der Tabelle 7.2

$(\sigma^2/95\%)$	direkt	N=48	N=198
x[cm]	(2232/90.5)	(296/ 32.4)	(99.3/18.6)
y[cm]	(2121/86.5)	(313/33.5)	(101/20.6)
Φ [°]	(180/35.6)	(24.2/9.8)	(8.0/5.5)
abs[cm]	(1502/125)	(116/56.3)	(26.7/44.1)

Tabelle 7.2. Statistische Kenngrößen für N=0, 48, 198

soll gezeigt werden, wie der „Masseneffekt“ die Ergebnisse im System verbessert. Hierfür wurden wieder Simulationen durchgeführt, diesmal allerdings mit 198 Hilfsknoten. Man kann sehen, wie die Varianz σ^2 bei größerer Helferzahl sinkt. Darüber hinaus wurde die 95%-Grenze in Tabelle 7.2 aufgenommen. Die 95%-Grenze gibt den Fehler an, bei der die kumulative Verteilung 0.95 erreicht hat; sie gibt also denjenigen höchsten Fehler an, der bei 95% aller Fälle eintreten kann.

Als nächster Schritt soll nun der Zusammenhang zwischen Knoten- oder Helferanzahl und der erreichten Genauigkeit betrachtet werden. Tabelle 7.2 gibt hier schon einen ersten Hinweis. Die Konvergenz für steigende N scheint naheliegend. Es wurde nun wieder die gleiche räumliche Situation simuliert. Die Anzahl der Knoten variierte in $N = [3..200]$. Für jede Knotenanzahl wurden 100 zufällige Topologien simuliert, wobei die Knoten A und B fest an den Positionen (150,150) und (350,350) mit zufälliger Orientierung verblieben. Aus diesen Simulationen wurden Mittelwert und Standardabweichungen ermittelt und in Abbildung 7.35 gegen die Anzahl von Knoten aufgetragen. Man kann sehen, wie die Standardabweichung sich verringert, je mehr Knoten involviert sind. Der Gradient ist besonders für kleine Anzahlen ($N < 10$) sehr hoch. Das bedeutet, dass sogar mit einer relativ kleinen Anzahl von Knoten eine erhebliche Verbesserung in der Genauigkeit erreicht werden kann. Da aber die Prozesszeit und der Aufwand, unabhängig von der Anzahl der Knoten, konstant ist, würde man natürlich stets alle möglichen Knoten miteinbeziehen.

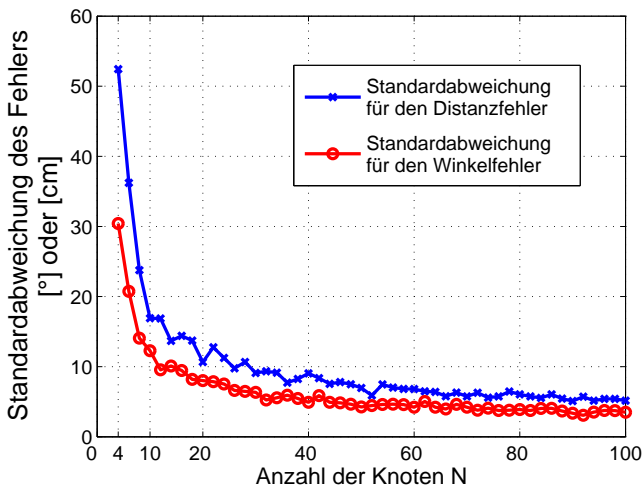


Abbildung 7.35. Graduelle Verbesserungen der Genauigkeit mit steigender Anzahl von Knoten N ; Die Unebenheiten in den Kurven sind statistische Streuungen.

7.5.12 Zusammenfassung

In den voranstehenden Unterkapiteln wurde ein Verfahren vorgestellt, wie man die Genauigkeit in Lokationssystemen verbessern kann. Das Verfahren hat geringe Anforderungen an die Hardware und erzeugt wenig Overhead. Die gesamte Prozesszeit ist unabhängig von der Anzahl der Teilnehmer. In der Analyse der bestehenden Hardwareplattform wurde erkannt, dass besonders die *systematischen* Fehler mit unbekanntem Modell kritisch für Lokationssysteme sind. In einem Lokationssystem kann ein Knoten A einen anderen Knoten B nicht präzise lokalisieren, wenn es große systematische Fehler im Messsystem gibt. Hier soll noch einmal das Fehlermodell aus (7.6) in Erinnerung gerufen werden: $\tilde{r} = r + f(\mathbf{t}, r) + n$. Der Teil des Fehlers, der als systematisch erkannt wurde ($f(\mathbf{t}, r)$) hängt vom Systemzustand \mathbf{t} ab. Für den Fall eines statischen Umfeldes wird der Systemstatus \mathbf{t} konstant sein und der Fehler somit wiederholbar. Dies konnte auch in den Experimenten in Abbildung 7.24 nachgewiesen werden. Wenn das Fehlermodell unbekannt, aber der Mittelwert Null ist, so kann der resultierende Fehler durch Mittelung verringert werden. Dies wurde mit Hilfe in Abschnitt 7.5.7 erklärt. Für Lokationssysteme wurde festgestellt, dass die systematischen Fehler am meisten von den Winkeln Φ, φ abhängen. Will man also den Systemzustand variieren, um die Mittelung nutzen zu können, so müsste man die gleiche topologische Konstellation unter verschiedenen Orientierungen messen. Da eine Rotation der Objekte nicht möglich ist, wurde in Abschnitt 7.5.8 vorgeschlagen, Messungen über zusammengesetzte Wege zu nehmen, um so verschiedene Ein- und Ausfallwinkel in den Endknoten zu erreichen.

Die Simulationen haben gezeigt, dass mit der Methode der mehrfachen Messwege die Genauigkeit der Lokalisierung von Objekten deutlich verbessert wird. Mit Hilfe von SDJS und der Rundrufnatur der Funkkommunikation wird so ein System erzeugt, das mit geringen Anforderungen an die Hardware trotzdem den Masseneffekt vieler, verteilter Objekte nutzen kann.

8 Zusammenfassung und Ausblick

Sensornetzwerke sind ein ideales Mittel, um physikalische Phänomene großflächig und feingranular zu beobachten. Die lokalen Sensorwerte werden aufgenommen, evtl. lokal vorverarbeitet, und dann durch das Netzwerk zur Informationssenkke transportiert. Die Sensorknoten haben hierbei eine Vielzahl von Prozessen abuarbeiten. Diese reichen von der eigentlichen Messung oder Beobachtung über Datenfusion und -speicherung bis hin zu komplexen Multihop-Protokollen und autonomer und adaptiver Netzwerkorganisation. Um solche Systeme effizient und stabil entwerfen zu können, müssen neue Mechanismen und Lösungen in allen Bereichen gefunden werden. Sensornetzwerke stellen neue Herausforderungen an die Skalierung von verteilten Prozessen, Kommunikation und Datenmanagement.

Ubiquitous Computing verfolgt die Vision der allgegenwärtigen Informationstechnologie. Es wird eine Vielzahl von Geräten in alltäglicher Umgebung vorhanden sein, die mit Sensorik und Aktuatoren ihre Umwelt erfassen und beeinflussen. Der Benutzer wird dabei sowohl über direkte Eingaben als auch über Sensorik erfasst und interagiert so mit den Systemen. Ubiquitous Computing skizziert eine Umgebung, in der Rechenvorgänge, Algorithmik, Datenauswertung und -speicherung stark verteilt werden. Um diese verteilten Vorgänge zu koordinieren, müssen die Einheiten der Informationstechnologie miteinander kommunizieren. Dies wird — aufgrund der Mobilität und Verteilung der Geräte — drahtlos geschehen.

8.1 Die Zielplattform

Die zwei Forschungsbereiche *Ubiquitous Computing* und *Sensornetzwerke* haben viele Gemeinsamkeiten. Die Systeme werden aus einer Vielzahl von Einheiten aufgebaut, die in drahtloser Kommunikation stehen. Die Datenprozesse reichen von der Messung über die Datenfusion bis hin zur Darstellung für den Benutzer. Aus den Forschungsbereichen *Ubiquitous Computing* und *Sensornetzwerke* wurde in der vorliegenden Arbeit die *Zielplattform* abgeleitet. Mit der Zielplattform

wird der Anwendungsbereich der präsentierten Beiträge festgelegt. Zur Erinnerung werden nun noch einmal kurz die Haupteigenschaften der Zielplattform genannt:

Die Zielplattform sind Geräte, die neben einer Funkschnittstelle nur sehr wenig Rechenleistung aufbringen. Die Funkschnittstelle charakterisiert sich dadurch, dass keine komplexe (De-) Modulation verarbeitet werden kann und ausschließlich Bandpasssignale zum Einsatz kommen. Die Rechenleistung der Zielplattform erlaubt keine hohen Abtastraten (über einige 10kHz) und kann keine aufwendige digitale Signalverarbeitung leisten.

Typische Vertreter der Klasse von Geräten, die durch die Zielplattform definiert wird, sind Sensorknoten, RFID, und sogar polymerelektronische Funklösungen. In der vorliegenden Arbeit werden Systeme betrachtet, die aus solchen Geräten aufgebaut sind. Das wichtigste Beispiel sind hier die Sensornetzwerke. In Kapitel 1.4 wurden vier zentrale Probleme in Informationssystemen wie Sensornetzwerken identifiziert. Diese sind

- Zeitsynchronisation,
- Kanalzugriff,
- Zuverlässigkeit,
- Datenfusion.

Diese Probleme werden mit bekannten Verfahren nur ungenügend gelöst. Daraus ergibt sich die Motivation der vorliegenden Arbeit, sie mit dem neuen Ansatz der *überlagerten Funksignale* zu lösen.

8.2 These und Beiträge

Die zentralen Beiträge der vorliegenden Arbeit werden in der These zusammengefasst und formuliert. Die damit aufgestellte Behauptung wurde mit den Einzelbeiträgen der Arbeit nachgewiesen.

These In leistungsschwachen Sensornetzwerken lösen überlagerte Funksignale grundsätzliche Probleme wie Synchronisierung, Zuverlässigkeit, Kanalnutzung und Datenfusion.

In Kapitel 3 werden zunächst allgemeine Modelle für eine Signalisierung mit überlagerten Funksignalen hergeleitet. Der wichtigste Beitrag ist hier das neue

Modulationsverfahren *ESK*, das überlagerte Funksignale für die leistungsschwache Ziellplattform ermöglicht. Darüber hinaus wird mit einem einfachen Bandpreizverfahren in Kapitel 3.1.3 die Detektion verbessert. Mit dem Einsatz von *Jamsignalen*, die eine einfache Darstellung einer solchen Bandspreizung sind, werden in Kapitel 4 die Themen *Zeitsynchronisation*, *Kanalzugriff* und *Zuverlässigkeit* bearbeitet. Die Lösungen werden im Protokoll AwareCon auf der pPart Particle Computer Plattform evaluiert. Die einzelnen Beiträge zur Bestätigung der These werden nun als Zusammenfassung präsentiert:

Zeitsynchronisation. AwareCon implementiert ein *Zeitsynchronisationsverfahren*, das auf überlagerten Funksignalen beruht. Dabei wird zusätzlich ein verteilter Mechanismus eingesetzt, der alle Knoten einer Zelle in gleicher Weise zufällig an der Synchronisation teilnehmen lässt. Alle Knoten sind in gleicher Weise dafür verantwortlich, die Synchronisation aufrecht zu erhalten. Mit diesem verteilten Algorithmus wird eine sehr schnelle Synchronisation erreicht. Die nötige Zeit, bis ein neuer Knoten mit einem Netz synchronisiert ist, ist typischerweise 12ms. Mit dem Austausch von Synchronisationssymbolen durch überlagerte Funksignale wird die Synchronisation dann aufrecht gehalten. Auf diese Weise wird eine Synchronisationsgenauigkeit von $4\mu\text{s}$ zwischen allen Netzwerkteilnehmern erreicht.

Kanalzugriff. Der *Kanalzugriff* in AwareCon nutzt ebenfalls überlagerte Funksignale. So ist es erstmals möglich, auf der Funkschnittstelle einen virtuellen tri-state-Bus zu realisieren. Dieser ist Voraussetzung für das Kanalzugriffsverfahren CSMA/CR, das so zum ersten Mal auch Anwendung in drahtloser Netzwerktechnik findet. Die besonderen Eigenschaften von CSMA/CR sind herausragende Skalierbarkeit, geringe Kollisionsraten und flexible Priorisierung.

Zuverlässigkeit. In Kapitel 6 wird das Motivationsbeispiel (das Sensor-Reach-Back-Problem) aufgegriffen und quantitativ analysiert. Es werden Szenarien diskutiert, in denen z.B. aufgrund von Installationsprozessen eine schlechte Konnektivität der Kommunikationsparteien besteht. Durch die schlechte Konnektivität ist auch die Zuverlässigkeit der diskutierten Systeme schlecht, da die Kommunikation zwischen Netzwerkpartnern teilweise nicht möglich ist. Um hier die Zuverlässigkeit von Verbindungen und die Konnektivität im Allgemeinen zu verbessern, werden überlagerte Funksignale eingesetzt. Die Idee ist hierbei, durch *kooperatives Senden* die Sendeleistung einer Gruppe von Funkknoten zu akkumulieren, um so die Reichweite zu erhöhen. So können z.B. unterbrochene Verbindungen wieder her-

gestellt und so die Zuverlässigkeit des Systems erhöht werden. In den Simulationen wird gezeigt, dass in bestimmten Konstellationen die Konnektivität im Vergleich zu herkömmlichen Multihop-Protokollen um über 50% verbessert werden kann.

Datenfusion. In Kapitel 7 wird SDJS als ein Protokoll für Datenprozesse in Sensornetzwerken vorgestellt. Mit SDJS und der Erweiterung Multi-SDJS sind vielfältige Datenprozesse möglich. Diese Datenprozesse haben die gemeinsame Eigenschaft, dass die zu verarbeitenden Daten stets in *verteilter* Form vorliegen. Mit SDJS können Kommunikation und Datenprozess in einem Vorgang vereint werden. Beispiele sind die Schätzung von Parametern oder Mittelwertbildung. Datenprozesse, die mit SDJS dargestellt werden, haben herausragende Skalierungseigenschaften. So ist z.B. die Prozessdauer unabhängig von der Anzahl der Teilnehmer, die Daten fusionieren wollen. In den Abschnitten 7.2 bis 7.5 wird SDJS und die Erweiterung Multi-SDJS für verschiedene Datenfusionsprozesse wie Anzahlsschätzung und Mittelwertbildung vorgestellt, implementiert und evaluiert. SDJS kann z.B. die Anzahl von Kommunikationsgeräten in einer Zelle bis zu 1000mal schneller schätzen, als dies mit klassischen Netzwerkverfahren möglich ist.

Mit der vorliegenden Arbeit werden *überlagerte Funksignale* als ganzheitliches Konzept für viele technischen Fragestellungen in Sensornetzwerken und Ubiquitous Computing verwendet. Überlagerte Funksignale sind eine innovative und unkonventionelle Lösung für die neuen Anforderungen in allgegenwärtiger Informationstechnologie. Die bekannten Verfahren mit überlagerten Funksignalen sind nur auf jeweils sehr spezielle Fragestellungen in Funknetzen oder Anwendungen zugeschnitten und erreichen auch keine allgemeine Anwendbarkeit auf der diskutierten Zielpattform. Speziell die Systemvoraussetzungen bekannter Lösungsvorschläge, basierend auf überlagerten Funksignalen, sind für Sensornetzwerke häufig nicht zu erfüllen.

Mit den Konzepten aus Kapitel 3 wurde ein Übertragungssystem vorgeschlagen, das auf *nicht-kohärent* überlagerten Funksignalen basiert, und durch seine geringen Anforderungen an die Hardware breite Einsatzgebiete weit über Sensornetzwerke hinaus ermöglicht. Andere, bekannte Lösungen ohne *nicht-kohärente* Überlagerungen haben nur sehr begrenzte Anwendungsbereiche.

Mit dem Einzug von überlagerten Funksignalen in die Kommunikationstechnik der drahtlosen Sensornetzwerke können für viele Bereiche der Kommunikation und Systemdienste völlig neue und effiziente Lösungen erbracht werden.

Tabelle 8.1 zeigt noch einmal die wichtigsten Beiträge der vorliegenden Arbeit in einer quantitativen Übersicht:

	Beitrag	Ergebnis
überlagerte Funksignale	nichtkohärente Verfahren	neuartige Mechanismen
Kanalzugriff	CSMA/CR für Funkssysteme	hohe Skalierbarkeit, geringe Kollisionsraten
Kanalzugriff	TOMAC	Echtzeitfähigkeit
Zeit-synchronisation	überlagerte Synchronisations-symbole	4 μ s genau Einbuchzeit 12ms
Zuverlässigkeit	kooperatives Senden	Steigerung der Zuverlässigkeit um 50%
Datenfusion	Parameterschätzung mit SDJS	Beschleunigung um Faktor 1000
Datenfusion	Datenverarbeitung mit Multi-SDJS	Beschleunigung um Faktor 60

Tabelle 8.1. Quantitative Erfassung der Beiträge dieser Arbeit

Jeder Einzelbeitrag der vorliegenden Arbeit zeigt für sich die herausragende Leistungsfähigkeit und Anwendbarkeit von überlagerten Funksignalen in Sensornetzwerken. Sämtliche Vorschläge und Lösungen, die in der vorliegenden Arbeit präsentiert wurden, sind auf der Sensornetzwerkplattform pPart Particle Computer implementiert und praktisch nachgewiesen worden. Viele der Mechanismen sind fester Bestandteil des Funkprotokolls AwareCon der pParts geworden und seit einigen Jahren weltweit im Einsatz.

8.3 Ausblick

Das Konzept der überlagerten Funksignale, so wie es in der vorliegenden Arbeit vorgestellt wird, besticht vor allem durch zwei Eigenschaften:

1. Die Vorteile von überlagerten Funksignalen können bereits mit sehr geringen technischen Voraussetzungen genutzt werden.

2. Die Verbesserungen, die überlagerte Funksignale zu den diskutierten Bereichen beitragen, sind überdurchschnittlich. Quantitative Verbesserungen liegen im Bereich von 50% (Verbesserung der Konnektivität) bis zu Faktor 1000 (Geschwindigkeitsvorteil bei der Schätzung von Anzahlen).

Diese vielversprechenden Eigenschaften motivieren, überlagerte Funksignale auch weiterhin — insbesondere für Sensornetzwerke — zu erforschen. Obwohl die vorliegende Arbeit viele Beiträge bereits mit Implementierungen und Evaluierungen abgeschlossen hat, sind viele Verbesserungen möglich. So ist z.B. der Einsatz von Bandspreizen und Rauschsignaturen in Kapitel 3.1.3 zur Verbesserung der Detektion sicherlich ein Thema, das noch für weitere Untersuchungen offen ist. Besonders die quantitativen Zusammenhänge zwischen Symboldauer und Bandbreite und die effiziente Generierung und Modulation der Signale auf einer Sensorplattform sind hier noch interessante Forschungsthemen. Auch auf Protokollebene sind noch viele Fragen offen. So ist die Koordination von kooperativem Senden im Hinblick auf klassische Netzwerkprobleme noch nicht geklärt. Dies wurde bereits in Kapitel 6.5 dargestellt. Auch SDJS als Datenfusionsprotokoll kann weiter verbessert werden. So könnte man über Leistungsregelungen und verbesserter Detektion der Jamsignale im Empfänger versuchen, auch Kollisionen zu erkennen und dies in die Schätzung (mit der bekannten Ungewissheit) einfließen zu lassen. Nicht zuletzt ist auch der verteilte Synchronisationsmechanismus aus Kapitel 5.3 noch offen für weitere Verbesserungen, um Effizienz und Geschwindigkeit zu erhöhen.

Überlagerte Funksignale wurden in der gegenwärtigen wissenschaftlichen Literatur noch nicht erschöpfend behandelt. Ähnlich den Beiträgen der vorliegenden Arbeit, wird die Kreativität der Forscher in diesem Bereich sicherlich noch viele, völlig neue Anwendungen aufdecken, die durch die effizienten und verblüffend einfachen Mechanismen von überlagerten Funksignalen dargestellt werden können.

Literaturverzeichnis

- [1] *The Lofar Project Site – Agriculture*. Website, accessed: 10/2005. – <http://www.lofar.org/p/Agriculture.htm>
- [2] *Magellan Technology Pty. Ltd.* Website, accessed: 9/2004. – <http://www.magtech.com.au/>
- [3] *Metro Group - Future Store Initiative*. Website, accessed: 10/2004. – <http://www.future-store.org>
- [4] *The Particle Computer Company*. Website, accessed: 10/2005. – <http://www.particle-computer.de>
- [5] *The RELATE Project Page*. Website, accessed: 04/2004. – <http://www.comp.lancs.ac.uk/~hwg/relate.html>
- [6] AARABI, Parham: The Fusion of Distributed Microphone Arrays for Sound Localization. In: *EURASIP Journal of Applied Signal Processing (Special Issue on Sensor Networks)* (2004), Nr. 4, S. 338–347
- [7] AARABI, Parham ; MUNGAMURU, Bob: Scene Reconstruction Using Distributed Microphones. In: *Proceedings of the IEEE Conference on Multimedia and Expo (ICME)*. Baltimore, Maryland, July 2003
- [8] ANANTHASUBRAMANIAM, Bharath ; MADHOW, Upamanyu: Virtual Radar Imaging for Sensor Networks. In: *Proceedings of the 3rd International Symposium on Information Processing in Sensor Networks (IPSN)*, 2004
- [9] ANCEAUME, Emmanuelle ; PUAUT, Isabelle: *A Taxonomy of clock Synchronization Algorithms*. July 1997. – research report IRISA
- [10] BAR-DAVID, Israel ; PLOTNIK, Eli ; ROM, Raphael: Limitations of the capacity of the M-user binary adder channel due to physical considerations. In: *IEEE Transaction on Information Theory* 40 (1994), Nr. 3, S. 662–673
- [11] BARRIAC, Gwen ; ; MUDUMBAL, Raghuraman ; MADHOW, Upamanyu: Distributed beamforming for information transfer in sensor networks. In: *Proceedings of the in Third International Workshop on Information Processing in Sensor Networks (IPSN)*. Berkeley, USA, April 2004

- [12] BEIGL, Michael ; KROHN, Albert ; ZIMMER, Tobias ; DECKER, Christian ; ROBINSON, Philip: AwareCon: Situation Aware Context Communication. In: *Proceedings of Ubicomp 2003*. Seattle, USA, Oktober 2003
- [13] BEIGL, Michael ; ZIMMER, Tobias ; KROHN, Albert ; DECKER, Christian ; ROBINSON, Philip: Smart-Its – Communication and Sensing Technology for UbiComp Environments / University of Karlsruhe. 2003 (2003/2). – ISSN 1432-7864
- [14] BROCH, Josh ; MALTZ, David A. ; JOHNSON, David B. ; HU, Yih-Chun ; JETCHEVA, Jorjeta: A Performance Comparison of Multi-Hop Wireless Ad Hoc Network Routing Protocols. In: *Mobile Computing and Networking*, 1998, S. 85–97
- [15] CHEN, Deqiang ; LANEMAN, Nicholas J.: Noncoherent Demodulation for Cooperative Diversity in Wireless Systems. In: *Proceedings of the IEEE Global Communication Conference (GLOBECOM)*. Dallas, TX, USA, November 2004
- [16] DAM, Tijs van ; LANGENDOEN, K.: An Adaptive Energy-Efficient MAC Protocol for Wireless Sensor Networks. In: *The First ACM Conference on Embedded Networked Sensor Systems (SenSys 2003)*. Los Angeles, USA, November 2003
- [17] DECKER, C. ; KROHN, A. ; BEIGL, M. ; ZIMMER, T.: The Particle Computer System. In: *IPSN Track on Sensor Platform, Tools and Design Methods for Networked Embedded Systems (SPOTS). Proceedings of the ACM/IEEE Fourth International Conference on Information Processing in Sensor Networks*. Los Angeles, USA, 2005
- [18] ELSON, Jeremy ; GIROD, Lewis ; ESTRIN, Deborah: Fine-Grained Network Time Synchronization using Reference Broadcasts. In: *Proceedings of the Fifth Symposium on Operating Systems Design and Implementation (OSDI)*. Boston, USA, Dec 2002
- [19] GREUNEN, Jana van ; RABAEAY, Jan: Lightweight Time Synchronization for Sensor Networks. In: *Second ACM International Workshop on Wireless Sensor Networks and Applications in conjunction with ACM MobiCom 2003*. San Diego, CA, USA, September 2003
- [20] HAZAS, Mike ; KRAY, Christian ; GELLERSEN, Hans ; AGBOTA, Henoc ; KORTUEM, Gerd ; KROHN, Albert: A Relative Positioning System for

- Co-Located Mobile Devices. In: *Proceedings of the 3rd International Conference on Mobile Systems, Applications, and Services (MobiSys)*. Seattle, USA, June 6-8 2005
- [21] HERMANN, Markus. *Evaluierung der Empfangsleistungen von überlagerten Funksignalen*. Studienarbeit am TecO, Universität Karlsruhe. 2005
- [22] HIGHTOWER, Jeffrey ; BORRIELLO, Gaetano: Location Systems for Ubiquitous Computing. In: *IEEE Computer* 34 (2001), August, Nr. 8, S. 57–66
- [23] HIGHTOWER, Jeffrey ; WANT, Roy ; BORRIELLO, Gaetano: SpotON: An Indoor 3D Location Sensing Technology Based on RF Signal Strength / University of Washington, Department of Computer Science and Engineering. Seattle, WA, February 2000 (00-02-02). – UW CSE
- [24] HILLENBRAND, Joerg ; WEISS, Timo: Calculation of Detection and False Alarm Probabilities in Spectrum Pooling Systems. In: *IEEE COMMUNICATIONS LETTERS* 9 (2005), April
- [25] HOFMANN-WELLENHOF, B. ; LICHTENEGGER, H. ; COLLINS, J. *GPS Theory and Praxis*. 1997
- [26] HONG, Yao-Win ; SCAGLIONE, Anna: Energy-Efficient Broadcasting with cooperative Transmission in Wireless Sensor Networks. In: *IEEE Transaction on Wireless Communication* . – under submission
- [27] HONG, Yao-Win ; SCAGLIONE, Anna: Distributed Change Detection in Large Scale Sensor Networks Through the Synchronization of the Pulse-Coupled Oscillators. In: *Proceeding of the 2004 International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. Montreal, Canada, May 2004
- [28] HONG, Y.W. ; SCAGLIONE, A.: Time Synchronization and Reach-back Communications with Pulse-Coupled Oscillators for UWB Wireless Ad Hoc Networks. In: *Proceedings of the IEEE Conference on Ultra Wideband Systems and Technologies (UWBST)*. Reston, USA, November 2003
- [29] HONG, Y.W. ; SCAGLIONE, A.: Distributed Change Detection in Large Scale Sensor Networks Through the Synchronization of the Pulse-Coupled Oscillators. In: *Proceeding of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. Montreal, Canada, Mai 2004
- [30] HONG, Y.W. ; SCAGLIONE, A.: A simple Method to Reach Detection Consensus in Massively Distributed Sensor Networks. In: *International symposium on Information Theory (ISIT)*, 2004

- [31] HU, An swol ; SERVETTO, Sergio: dFSK: Distributed Frequency Shift Keying Modulation in Dense Sensor Networks. In: *Proceedings of the IEEE International Conference on Communications (ICC)*. Paris, France, June 2004
- [32] HU, An swol ; SERVETTO, Sergio: On the Scalability of Cooperative Time Synchronization in Pulse-Connected Networks. In: *IEEE Transactions on Information Theory* (2006), June
- [33] HU, An swol ; SERVETTO, Sergio D.: Algorithmic Aspects of the Time Synchronization Problem in Large-Scale Sensor Networks. In: *ACM/Kluwer Journal on Mobile Networks and Applications (MONET); Special Issue on Wireless Sensor Networks, with selected (and revised) papers from ACM WSNA 2003. Invited Paper*. 10 (2005), S. 491–503
- [34] HU, A.S ; SERVETTO, S: On the Scalability of cooperative Time Synchronization in Pulse-Connected Networks. In: *IEEE Transactions on Information Theory* to appear (2006)
- [35] IEEE COMPUTER SOCIETY, LAN/MAN Standards C. t.: Supplement to IEEE standard for information technology telecommunications and information exchange between systems - local and metropolitan area networks - specific requirements. Part 11: wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: high-speed physical layer in the 5 GHz band. In: *IEEE Std 802.11a-1999* (1999)
- [36] III, D. Richard B. ; PRINCE, Gregory B. ; MCNEILL, John A.: Method for Carrier Frequency and Phase Synchronization of two Autonomous Cooperative Transmitters. In: *5th IEEE Signal Processing Advances in Wireless Communications (SPAWC) 2005*, 2005
- [37] ISO STANDARD COMMITTEE: Road vehicles – Controller area network (CAN) – Part 1: Data link layer and physical signalling. (2003). – ISO 11898-1
- [38] ISO/IEC/JTC 1/ SC 31/WG 4: Automatic Identification - Radio Frequency Identification for Item Management Communications and Interfaces Part 3: Physical Layer, Anti collision System and Protocol Values at 13.56 MHz MODE 2. (2001)
- [39] KAUTZ, William H. ; SINGLETON, Roy. C.: Nonrandom Binary Superimposed Codes. In: *IEEE Transaction on Information theory* (1964)

- [40] KROHN, Albert: Optimal non-coherent m-ary energy shift keying for cooperative transmission in sensor networks. In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Toulouse, France, May 14-19 2006
- [41] KROHN, Albert ; BEIGL, Michael ; DECKER, Christian ; RIEDEL, Till ; ZIMMER, Tobias: The Implementation of Non-Coherent Cooperative Transmission for Wireless Sensor Networks. In: *Adjunct Proceedings of the International Conference on Pervasive Computing (PERVASIVE)*. Dublin, Ireland, Februar 2006
- [42] KROHN, Albert ; BEIGL, Michael ; DECKER, Christian ; RIEDEL, Till ; ZIMMER, Tobias: The Implementation of Non-Coherent Cooperative Transmission for Wireless Sensor Networks. In: *Proceedings of the International Conference on Networked Sensing Systems (INSS)*. Chicago, USA, Mai 2006
- [43] KROHN, Albert ; BEIGL, Michael ; DECKER, Christian ; ZIMMER, Tobias: TOMAC- Real-Time Message Ordering in Wireless Sensor Networks Using the MAC Layer. In: *2nd International Workshop on Networked Sensing Systems (INSS)*. San Diego, USA, June 2005
- [44] KROHN, Albert ; BEIGL, Michael ; HAZAS, Mike ; GELLERSEN, Hans ; SCHMIDT, Albrecht: Using Fine-Grained Infrared Positioning to Support the Surface-Based Activities of Mobile Users. In: *Proceedings of the 5th International Workshop on Smart Appliances and Wearable Computing (IWSAWC)*. Columbus, USA, 2005
- [45] KROHN, Albert ; BEIGL, Michael ; WENDHACK, Sabin: SDJS: Efficient Statistics for Wireless Networks. In: *Proceedings of the 12th IEEE International Conference on Network Protocols*. Berlin, Germany, 2004
- [46] KROHN, Albert ; ZIMMER, Tobias ; BEIGL, Michael ; DECKER, Christian: Collaborative Sensing in a Retail Store Using Synchronous Distributed Jam Signalling. In: *3rd International Conference on Pervasive Computing*. Munich, Germany, 2005
- [47] KRUMM, John ; HORVITZ, Eric: LOCADIO: Inferring Motion and Location from Wi-Fi Signal Strengths. In: *Proceedings of the First Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services (MobiQuitous)*. Boston, USA, August 2004, S. 4–13

- [48] LANEMAN, J. N. ; TSE, David N. C. ; WORNELL, Gregory W.: Cooperative Diversity in Wireless Networks: Efficient Protocols and Outage Behavior. In: *IEEE Transactin on information Theory* 50 (2004), December, Nr. 12
- [49] LANGENDOEN, Koen ; REIJERS, Niels: Distributed Localization in Wireless Sensor Networks: A Quantitative Comparison. In: *Computer Networks* 43 (2003), August, S. 499–518
- [50] LAN/MAN STANDARDS COMMITTEE OF THE IEEE COMPUTER SOCIETY: Information technology - telecommunications and information exchange between systems - local and metropolitan area networks - specific requirements. Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) access method and physical layer specifications. In: *IEEE Std. 802.3-1998* (1998)
- [51] LAN/MAN STANDARDS COMMITTEE OF THE IEEE COMPUTER SOCIETY: IEEE Standard for Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Specific requirements Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (LR-WPANs). In: *IEEE Std 802.15.4-2003* (2003)
- [52] LAVAL, Peter de ; ABDUL-JABBAR, Shakir: Decoding of superimposed codes in multiaccess communication. In: *8th European Conference on Area Communication, EUROCON*, 1988
- [53] MAHAJAN, Ajay ; WALWORTH, Maurice: 3D position sensing using the differences in the time-of-flights from a wave source to various receivers. In: *IEEE Transaction on Robotics and Automation* 17 (2001), february, Nr. 91
- [54] MARÓTI, Miklós ; KUSÝ, Branislav ; BALOGH, György ; VÖLGYESI, Péter ; NÁDAS, András ; MOLNÁR, Károly ; DÓRA, Sebestyén ; LÉDECZI Ákos: Radio Interferometric Geolocation. In: *Proceedings of the Third International Conference on Embedded Networked Sensor Systems (SenSys)*. San Diego, USA, November 2005, S. 1–12
- [55] MARÓTI, Miklós ; KUSÝ, Branislav ; SIMON, Gyula ; LÉDECZI Ákos: The Flooding Time Synchronization Protocol. In: *Proceedings of the Second ACM Conference on Embedded Networked Sensor Systems (SenSys)*. Baltimore, USA, Nov. 2004

- [56] METRO GROUP FUTURE STORE INITIATIVE. *RFID: Uncovering the Value - Applying RFID within the Retail and Consumer Goods Value Chain*. White Paper, Januar 2004
- [57] MILLS, David: *Network Time Protocol - Version 3 Specification, Implementation and Analysis (rfc 1305)*. – <http://www.faqs.org/ftp/rfc/rfc1305.pdf>
- [58] MINAMI, Masateru ; FUKUJU, Yasuhiro ; HIRASAWA, Kazuki ; YOKOYAMA, Shigeaki ; MIZUMACHI, Moriyuki ; MORIKAWA, Hiroyuki ; AOYAMA, Tomonori: DOLPHIN: A Practical Approach for Implementing a Fully Distributed Indoor Ultrasonic Positioning System. In: *Proceedings of the Sixth International Conference on Ubiquitous Computing (UbiComp)*. Nottingham, UK : Springer, September 2004, S. 347–365
- [59] MOORE, David ; LEONARD, John ; RUS, Daniela ; TELLER, Seth: Robust Distributed Network Localization with Noisy Range Measurements. In: *Proceedings of the Second International Conference on Embedded Networked Sensor Systems (SenSys)*. Baltimore, USA, November 2004, S. 50–61
- [60] MUNGAMURU, Bob ; AARABI, Parham: Joint Sound Localization and Orientation Estimation. In: *Proceedings of the 6th Int'l Conf. on Information Fusion (ICIF)*. Cairns, Australia, July 2003
- [61] NICULESCU, Dragoş ; NATH, Badri: Ad Hoc Positioning System (APS) using AOA. In: *Proceedings of the Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*. San Francisco, USA, März 2003, S. 1734–1743
- [62] NOSRATINIA, Aria ; HUNTER, Todd E. ; HEDAYAT, Ahmadrza: Cooperative Communication in Wireless Networks. In: (*accepted for publication in IEEE Communications Magazine*)
- [63] PRIYANTHA, Nissanka B. ; MIU, Allen K. L. ; BALAKRISHNAN, Hari ; TELLER, Seth: The Cricket Compass for Context-Aware Mobile Applications. In: *Proceedings of the Seventh International Conference on Mobile Computing and Networking (MobiCom)*. Rome, Italy, Juli 2001
- [64] PUGH, Jim ; MARTINOLI, Alcherio: Relative Localization and Communication Module for Small-Scale Multi-Robot Systems. In: *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2006

- [65] RFID JOURNAL: *Magellan unveils tunnel reader*. Website, accessed 9/2004. – <http://www.rfidjournal.com/article/articleview/445/1/1/>
- [66] RIBEIRO, Alejandro ; CAI, Xiaodong ; GIANNAKIS, Georgios B.: Symbol Error Probabilities for General Cooperative Links. In: *Proceedings fo the IEEE International Conference on communications* Bd. 6, 2004, S. 3369–3373
- [67] RINGWALD, Matthias ; RÖMER, Kay: BitMac: A Deterministic, Collision-Free, and Robust MAC Protocol for Sensor Networks. In: *roceedings of 2nd European Workshop on Wireless Sensor Networks (EWSN 2005)*. Istanbul, Turkey, january 2005, S. 57–69
- [68] RUPPRECHT, Werner: *Signale und Übertragungssysteme - Modelle und Verfahren*. Springer, 1993
- [69] SAVVIDES, Andreas ; GARBER, Wendy L. ; MOSES, Randolph L. ; SRIVASTAVA, Mani B.: An Analysis of Error Inducing Parameters in Multihop Sensor Node Localization. In: *IEEE Transactions on Mobile Computing* 4 (2005), November, Nr. 6, S. 567–577
- [70] SAVVIDES, Andreas ; HAN, Chih-Chieh ; SRIVASTAVA, Mani B.: Dynamic Fine-Grained Localization in Ad-Hoc Networks of Sensors. In: *Proceedings of the Seventh International Conference on Mobile Computing and Networking (MobiCom)*. Rome, Italy, Juli 2001, S. 166–179
- [71] SCAGLIONE, A. ; HONG, Y.-W.: Opportunistic Large Arrays: Cooperative Transmission in Wireless Multihop Ad Hoc Networks to Reach far Distances. In: *IEEE Transaction on Signal Processing* 51 (2003), August, Nr. 8
- [72] SHANG, Yi ; RUML, Wheeler: Improved MDS-Based Localization. In: *Proceedings of the 23rd Conference of the IEEE Communications Society (INFOCOM)*. Hong Kong, März 2004
- [73] SICHITIU, Mihail L. ; VEERARITTIPHAN, Chanchai: Simple, Accurate Time Synchronization for Wireless Sensor Networks. In: *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*. New Orleans, USA, März 2003
- [74] SIRKECI-MERGEN, B. ; SCAGLIONE, A.: SIGNAL ACQUISITION FOR COOPERATIVE TRANSMISSIONS IN MULTI-HOP AD-HOC NET-

- WORKS. In: *Proceeding of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. Montreal, Canada, May 2004
- [75] SIRKECI-MERGEN, Birsen ; SCAGLIONE, Anna: Coverage Analysis of Cooperative Broadcast in Wireless Networks. In: *Proceedings of the Fifth IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. Lisbon, Portugal, July 11-14 2004
- [76] STRAIGHT, Joseph H.: *Combinatorics: an Invitation*. Belmont, USA : Cole Publishing Company, 1993
- [77] THELEN, John ; GOENSE, Dann ; LANGENDOEN, Koen: Radio Wave Propagation in Potato Fields. In: *1st Workshop on Wireless Network Measurements*. Trentino, Italy, April 2005
- [78] TU, Yung-Szu ; POTTIE, Gregory: Coherent Cooperative Transmission From Multiple Adjacent Antennas To a Distant Stationary Antenna Through AWGN Channels. In: *Proceeding of the IEEE VTC Spring 02*. Birmingham, Alabama, USA, May 2002
- [79] VALIN, Jean-Marc ; MICHAUD, Francois ; ROUAT, Jean ; LETOURNEAU, Dominic: Robust Sound Source Localization Using a Microphone Array on a Mobile Robot. In: *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2003
- [80] ČAPKUN, Srdjan ; HAMDI, Maher ; HUBAUX, Jean-Pierre: GPS-free Positioning in Mobile Ad-Hoc Networks. In: *Cluster Computing Journal* 5 (2002), April, Nr. 2, S. 157–167
- [81] VOGT, Harald: Multiple Object Identification with Passive RFID Tags. In: *IEEE International Conference on Systems, Man and Cybernetics (SMC '02)*, 2002
- [82] WEISER, Mark: The computer of the 21st century. In: *Scientific American* 265 (1991), September, Nr. 3, S. 94–104
- [83] WEISS, Timo ; HILLENBRAND, Joerg ; JONDRAL, Friedrich: A Diversity Approach for the Detection of Idle Spectral Resources in Spectrum Pooling Systems. In: *48. Internationales Wissenschaftliches Kolloquium, TU Ilmenau*. TU Ilmenau, Germany, 2003, S. 37–38
- [84] WHITEHOUSE, Kamin ; KARLOF, Chris ; WOO, Alec ; JIANG, Fred ; CULLER, David: The Effects of Ranging Noise on Multihop Localization: An Empirical Study. In: *Proceedings of the Fourth International Symposium*

- on Information Processing in Sensor Networks (IPSN)*. Los Angeles, USA, April 2005, S. 73–80
- [85] YE, Wei ; HEIDEMANN, John ; ESTRIN, Deborah. *An Energy-Efficient MAC Protocol for Wireless Sensor Networks*. June 2002
- [86] YEDAVALLI, Kiran ; KRISHNAMACHARI, Bhaskar ; RAVULA, Sharmila ; SRINIVASAN, Bhaskar: Ecolocation: A Sequence Based Technique for RF Localization in Wireless Sensor Networks. In: *Proceedings of the Fourth International Symposium on Information Processing in Sensor Networks (IPSN)*. Los Angeles, USA, April 2005, S. 285–292

Abbildungsverzeichnis

1.1	Ein Funkknoten ist die kleinste Einheit eines Sensornetzwerkes, hier im Vergleich zur Blüte eines Gänseblümchens	3
1.2	Das Sensor-Reach-Back-Problem	4
1.3	Die Probleme, die aus dem Arbeitsbereich resultieren, motivieren zu näheren wissenschaftlichen Betrachtungen	7
1.4	Aufbau der Arbeit	8
2.1	Überblick über verschiedene Arten von überlagerten Funksignalen	13
2.2	Kohärente Überlagerung führt zu konstruktiver Interferenz	15
2.3	Inkohärente Überlagerung führt zu Schwebungen	15
2.4	Die Position der Knoten im Beispiel der Opportunistic Large Arrays, aus [71, S. 2085]	20
2.5	Der zeitliche Verlauf der Signalisierung für Opportunistic Large Arrays, aus [71, S. 2085]	21
2.6	Verstärkung von Signalen für die verteilte Detektion; aus [83] . . .	23
2.7	dFSK: Ergebnis der Überlagerung; aus [31, Abbildung 2]	25
2.8	Kohärent überlagerte Signale; aus [78, Abbildung 1, Seite 2] . . .	26
2.9	Kohärente Überlagerung von Signalen zur Erzeugung einer Richtcharakteristik; aus [36, Abbildung 8, Seite 8]	28
2.10	Ein mobiler Funkknoten sammelt Funksignale aus einem Sensornetz ein. Das System imitiert eine Radarstrahlreflektion; aus [8, Abbildung 1, Seite 2]	30
2.11	Relative Zeitsynchronisation mit <i>RBS</i> in einer Multihop-Topologie; aus [18, Abbildung 7a, Seite 14]	34
2.12	Ein Paket wird von Knoten 1 nach 2 geschickt und direkt wieder zurückgeschickt. Die Zeitpunkte t_0 , t_b und t_r werden notiert.; aus [73, Abbildung 1, Seite 3]	35
2.13	Die Amplitude des Empfangssignals in weiter Entfernung normalisiert auf die Leistung nur eines Senders; aus [28, Abbildung 3] .	36

2.14	Die Überlagerung von 400 Rechtecksignalen (unten). Sie stammen von 400 Knoten, die eine leichte Zeitverschiebung haben, die symmetrisch und offsetfrei um die Flanke des Rechtecks verteilt sind; aus [34, Abbildung 1]	37
2.15	Überlagerung punktsymmetrischer Impulse. Der Nulldurchgang ist der Synchronisationszeitpunkt; aus [34, Abbildung 5]	39
2.16	Thematische Einordnung dieser Arbeit: Es geht um Funksignale, die nicht orthogonal, kontrolliert und nicht kohärent im Bandpassbereich überlagert werden und nur skalar verarbeitet werden	42
2.17	Architektur der Kommunikation und der Transceiver der Sensor-knoten für überlagerte Funksignale	44
3.1	Systemmodell für überlagerte Funksignale	48
3.2	Kohärente Überlagerung führt zu konstruktiver Interferenz	49
3.3	Inkohärente Überlagerung führt zu Schwebungen	49
3.4	Rotation und Überlagerung im nicht-kohärenten Empfänger	51
3.5	Die tatsächliche Verteilung für $N = 8$ ist der Gaußverteilung sehr ähnlich	54
3.6	Der Kullback-Leibler-Abstand zwischen der tatsächlichen Verteilung auf I-Achse und der Gaußverteilung bzw. zwischen der tatsächlichen Amplitudenverteilung und der Rayleighverteilung	55
3.7	Verteilungen der Empfangleistung zweier überlagerter Signale: Bei Verletzung des Abtasttheorems ist die Verteilung nicht mehr zur Detektion geeignet.	58
3.8	Unterschiedliche Systemmodelle für kooperatives und verteiltes Senden	60
3.9	Optimale Entscheidungsgrenze $\gamma_{1 \rightarrow 2}$	64
3.10	Signalkonstellation in der komplexen Signalebene: Vergleich traditioneller ASK und PSK mit ESK für überlagerte Funksignale. Die Signalkonstellation für ESK ist rotationsinvariant – die Signalepunkte können auf Kreisen beliebig platziert werden.	68
4.1	Überlagerung von Funksignalen	72
4.2	Transformation der Messwerte von einer logarithmischen (links) in eine lineare (rechts) Skala; aus [21, Abbildung 18]	74

4.3	Ermittelte Energiemaße bei verschiedenen Versuchen; aus [21, Abbildung 20]. Die Versuche sind aufsteigend sortiert nach der mittleren empfangenen Energie. Diese Energie hängt z.B. von der geometrischen Position der Knoten ab, die in den Versuchen unterschiedlich war.	74
4.4	Confusion-Matrix für die Erkennung von mehrfach überlagerten Funksignalen	75
4.5	Versuchsaufbau zu Sensor-Reach-Back-Problem; aus [41]	76
5.1	Ein pPart Particle Computer Sensorknoten	79
5.2	Der Zeitrahmen des AwareCon-Protokolls	81
5.3	Kanalzugriff im AwareCon-Protokoll mit CSMA/CR. Aus [17, Abbildung 6]	82
5.4	Identische Aussendungen von zwei Sendern mit geringen Zeitsynchronisationsungenauigkeiten. Der Empfänger wird leicht gestreckte 1-Bits und leicht verkürzte 0-Bits empfangen (aus [67, Abbildung 5, Seite 6]).	83
5.5	Der Ablauf der Synchronisation im AwareCon-Protokoll	86
5.6	Der Synchronisationsmechanismus des AwareCon-Protokolls mit überlagerten Funksignalen	87
5.7	Die Verzögerungszeit, bis Teilnehmer im AwareCon-Protokoll synchronisiert sind	88
5.8	Die Darstellung der „non-destructive bit-wise arbitration“	91
5.9	Die Priorität einer wartenden Sendung nimmt mit zunehmender Wartezeit zu.	96
6.1	Erhöhung der Funkreichweite durch Summation der Sendeleistungen	102
6.2	Das Kommunikationsszenario: Knoten 0 (Quelle) will eine Nachricht an Knoten 6 (Ziel) verschicken	105
6.3	Das Kommunikationsszenario mit kooperativem Senden und Wellenausbreitung	106
6.4	Das Kommunikationsszenario und akkumulierendes, kooperatives Senden	107
6.5	Das Kommunikationsszenario mit akkumulierendem, kooperativem Senden	108
6.6	Das Kommunikationsszenario mit dem <i>hybriden</i> Verfahren	109
6.7	Das Kommunikationsszenario mit dem <i>hybriden</i> Verfahren	109

6.8	Konnektivität der Knoten mit ihrer Basisstation, Fläche=500m x 500m, Funkreichweite ≈ 50 m	111
6.9	Peer-to-Peer-Konnektivität von Knoten	112
6.10	Anzahl der Knoten mit Verbindung zur Basisstation	113
6.11	Grenzgewinn mit kooperativem Senden	115
6.12	Zwei Gruppen von Sensorknoten, die nur über kooperatives Senden miteinander kommunizieren können.	119
7.1	Architektur der IEEE 802.x Standards	122
7.2	Die Slotsstruktur des SDJS Verfahrens	125
7.3	Wahrscheinlichkeitsverteilung für Jamsignale; $s=64$	129
7.4	ML- und MAP Schätzer für $s = 32$	130
7.5	Invarianz des MAP/ML-Schätzers	131
7.6	SDJS-Simulation mit $s = 256$	132
7.7	Ein Experiment mit über 40 Particle Computern	133
7.8	Darstellung der Schätzergebnisse des Experiments mit 40 Particle Computern	133
7.9	Akkumulierte Fehlerverteilung im Experiment mit 40 Particle Computern; Vergleich mit der Simulation	134
7.10	Einige Zahlen über den manuellen Überprüfungsvorgang. Aus [46, Abbildung 2]	137
7.11	Profil von Mindesthaltbarkeitsdaten	138
7.12	Ein Multi-SDJS-Verfahren mit mehreren SDJS-Prozessen	140
7.13	Wahrscheinlichkeitsdichte für $k = 200, s = 200, 400, 800$	141
7.14	Akkumulierte Wahrscheinlichkeitsverteilung für den Schätzfehler bei $k=200$ Tags	142
7.15	Profil der Mindesthaltbarkeitsdaten für 50 Produkte beim Experiment mit der Implementierung	144
7.16	Einsammeln und Mittelung mit dem Multi-SDJS-Prozess: mehrere Sensorknoten mitteln ihre Temperaturmessung	146
7.17	Der Multi-SDJS-Prozess für das Einsammeln von Daten	147
7.18	Relative Positionierung von Knoten	152
7.19	Ein „Relateobjekt“ ist ein mobiler Sensorknoten für Ultraschallmessungen	154
7.20	Genauigkeit der rohen Messwerte für die Ultraschall-Plattform aus [20, Abbildung 11]	155
7.21	Verteilung der Empfindlichkeit der Schallwandler in Abhängigkeit vom Einfallswinkel φ	155

7.22	Messfehlerkurve für den Empfangswinkel	157
7.23	Mittlerer Fehler für den Empfangswinkel bei verschiedenen Entfernungen	158
7.24	Wiederholen von Empfangswinkelmessungen bei gleichem Systemzustand t . Die Fehlermaße repräsentieren die 5%- und 95%-Bereiche. Die Streuung ist relativ klein — dies zeigt, dass die Fehler tatsächlich wiederholbar identisch, also systematischer Natur sind.	159
7.25	Die Wahrscheinlichkeitsdichte für das Rauschen auf dem Empfangswinkel	159
7.26	Der statistische Fehler in der Abstandsmessung, verursacht durch Quantisierung und Rauschen	160
7.27	Der mittlere Fehler der Abstandsmessung	161
7.28	Die Verbesserungen in der Genauigkeit, wenn mehrere Messwerte gemittelt werden; die Abbildung zeigt kumulative Wahrscheinlichkeitsdichten für verschiedene Anzahlen von Messungen, die gemittelt werden.	163
7.29	Die 95% Fehlerkurve in Abhängigkeit von der Anzahl der gemittelten Messwerte: die Mittelung von nur wenigen Messwerten erzeugt schon eine relevante Verbesserung.	164
7.30	Mehrere Messpfade führen von Knoten A nach B	165
7.31	Verteilte Messwerte werden mit dem Multi-SDJS-Prozess eingesammelt	168
7.32	Ein Beispiel aus der Simulation für eine Konstellation mit 50 Knoten	170
7.33	Die Messungen können durch Mittelung verbessert werden: ihre Varianz wird kleiner, die Dichtekurven werden zusammengescho-ben	171
7.34	1000 Simulationen mit 50 Knoten	172
7.35	Graduelle Verbesserungen der Genauigkeit mit steigender Anzahl von Knoten N ; Die Unebenheiten in den Kurven sind statistische Streuungen.	174
B.1	Der Ablauf des TDMA-Protokolls AwareCon	201

Tabellenverzeichnis

2.1	Charakterisierung anderer Arbeiten auf dem Gebiet der überlagerten Funksignale	40
3.1	Äquivalente Ausdrücke in der Theorie zum Rayleigh-Kanal und nicht-kohärenter Überlagerung	53
6.1	Simulationsparameter	111
6.2	Vergleich der Kommunikationsprinzipien	117
7.1	Statistische Charakteristika für $N=0, 48$	173
7.2	Statistische Kenngrößen für $N=0, 48, 198$	173
8.1	Quantitative Erfassung der Beiträge dieser Arbeit	180

Glossar

Drahtlose Sensornetzwerke: Drahtlose Sensornetzwerke sind Systeme, die aus einer Vielzahl von drahtlos miteinander kommunizierenden Kleinstrechnern zusammengesetzt sind. Diese Kleinstrechner übernehmen z.B. Messaufgaben (z.B. von physikalischen Parametern) und kommunizieren sie über eine Funkschnittstelle, Seite 2.

Funkknoten: siehe Sensorknoten, Seite 2.

Jamsignale: Jamsignale sind zufällige, bandbegrenzte Zeitsignale. Jamsignale werden nur durch ihre Leistung, die Bandbreite und die Signaldauer festgelegt, Seite 93.

Sensorknoten: Sensorknoten sind die kleinste Einheit von Sensornetzwerken. Es sind Kleinstrechner mit Sensorik und Funkkommunikationsschnittstelle, Seite 2.

Ubiquitous Computing: Ein Wissenschaftszweig, der sich mit allgegenwärtiger Informationstechnologie befasst, Seite 1.

Überlagerte Funksignale: Überlagerte Funksignale sind Signale im Bandpassbereich, die zeitgleich ausgesendet werden und sich so im Funkkanal nicht orthogonal und inkohärent überlagern, Seite 41.

A Signalkonstellation

Um das globale Maximum von (3.31) zu finden, wird substituiert:

$$a = \sqrt{\frac{\sigma_{i+1}^2}{\sigma_{i-1}^2}} \quad \text{der zufällige Parameter} \quad (\text{A.1})$$

$$x = \sqrt{\frac{\sigma_i^2 \sigma_i^2}{\sigma_{i-1}^2 \sigma_{i+1}^2}} \quad \text{die neue Veränderliche} \quad (\text{A.2})$$

Eine generelle Bedingung ist, dass $\sigma_{i-1}^2 < \sigma_i^2 < \sigma_{i+1}^2$ ist und mit (A.1) und (A.2) sieht man, dass $\frac{1}{a} < x < a$ ist. Mit $\sigma_{i+1}^2 \geq \sigma_i^2$ folgt weiter $a > 1$. Mit (A.1), (A.2) in (3.32) wird definiert:

$$g_a(x) := -e^{\frac{\ln(ax)}{1-ax}} + e^{\frac{\ln(ax)}{1-ax}} - e^{\frac{\ln(\frac{a}{x})}{a-1}} + e^{\frac{\ln(\frac{a}{x})}{1-\frac{a}{x}}} \quad (\text{A.3})$$

$$\begin{aligned} f(y) &:= y^{\frac{1}{1-y}} - \frac{1}{y} \cdot y^{\frac{1}{1-y}} \\ \Rightarrow g_a(x) &= f(ax) + f\left(\frac{a}{x}\right) \end{aligned} \quad (\text{A.4})$$

$$q(y) := \frac{\partial f(y)}{\partial y} \cdot y = y^{\frac{-1}{y-1}} \frac{\ln y}{y-1} \quad (\text{A.5})$$

so dass

$$\frac{d f(ax)}{d x} = q(ax) \frac{1}{x} \quad \text{und} \quad \frac{d f(\frac{a}{x})}{d x} = -q\left(\frac{a}{x}\right) \frac{1}{x} \quad (\text{A.6})$$

und

$$\frac{d g_a(x)}{d x} = \frac{1}{x} \left(q(ax) - q\left(\frac{a}{x}\right) \right) \quad (\text{A.7})$$

Es kann gezeigt werden, dass $q(y)$ für $y > 1$ eine streng monoton sinkende Funktion ist:

$$q'(y) = \underbrace{y^{\frac{-1}{y-1}}}_{>0} \frac{\overbrace{1}^{>0}}{y(y-1)^3} \underbrace{\left((y-1)^2 - (y^2-1) \ln y + y \ln^2 y \right)}_{=:r(y)} \quad (\text{A.8})$$

und $r(y) < 0$ für $y > 1$, da

$$r(1) = \frac{d}{dy} r(1) = \frac{d^2}{dy^2} r(1) = 0$$

und

$$\frac{d^3}{dy^3} r = \underbrace{\frac{-2}{y^3}}_{<0} \underbrace{(y^2 - 1 + y \ln y)}_{=:s(y)}$$

und $s(1) = 0$, $s'(y) = 2y + \ln y + 1 > 0$ für $y > 1$.

Da nun $q'(y) < 0$ für $y > 1$ gilt, kann man zurückblickend auf (A.6) und (A.7) sehen, dass gilt:

$$\begin{aligned} x > 1 & : ax > \frac{a}{x} \Rightarrow q(ax) < q\left(\frac{a}{x}\right) \Rightarrow \frac{d}{dx} g_a < 0 \\ x < 1 & : ax < \frac{a}{x} \Rightarrow q(ax) > q\left(\frac{a}{x}\right) \Rightarrow \frac{d}{dx} g_a > 0 \\ x = 1 & : ax = \frac{a}{x} \Rightarrow \frac{d}{dx} g_a = 0 \end{aligned}$$

Dies bedeutet, dass $x = 1$ das einzige *globale* Maximum auf $g_a(x)$ ist. Einsetzen von $x = 1$ in (A.2) führt zu:

$$\frac{\sigma_{i+1}^2}{\sigma_i^2} = \frac{\sigma_i^2}{\sigma_{i-1}^2} \quad (\text{A.9})$$

Beachtet man, dass $\sigma_i^2 > \sigma_{i-1}^2$ ist, so kann man (3.33) als Erfüllung von (A.9) erkennen.

B Das AwareCon-Protokoll: Flussdiagramm

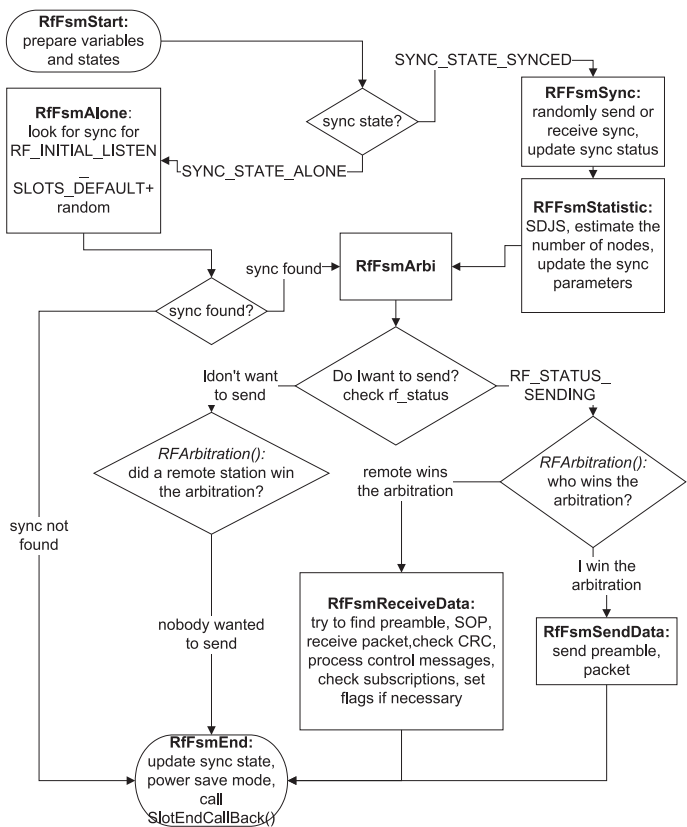


Abbildung B.1. Der Ablauf des TDMA-Protokolls AwareCon